Access DB# _____

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner # :_____ Date: _____
Art Unit: _____ Phone Number 30_____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher _____
Searcher Phone #: _____
Searcher Location: _____
Date Searcher Picked Up: _____
Date Completed: 2/3/01
Searcher Prep & Review Time: _____
Clerical Prep Time _____
Online Time: _____

Type of Search

NA Sequence (#)_____
AA Sequence (#)_____
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbit _____
Dr.Link _____
Lexis/Nexis_____
Sequence Systems _____
WWW/Internet _____
Other (specify)_____

PTO-1590 (1-2000)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 01:51:23 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	296	81.3	226	12	O09727	O09727 leonurus mo
2	296	81.3	226	12	Q9WHF6	Q9whf6 tomato mild
3	294	80.8	361	12	Q67574	Q67574 bean golden
4	292	80.2	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	292	80.2	364	12	Q9QS55	Q9qs55 sweet potat
6	289	79.4	185	12	Q98693	Q98693 sida golden
7	281	77.2	149	12	P88975	P88975 macroptiliu
8	281	77.2	233	12	Q9YLA4	Q9yla4 macroptiliu
9	275	75.5	234	12	O39180	O39180 potato yell
10	272	74.7	190	12	Q9Z089	Q9z089 tobacco lea
11	272	74.7	190	12	Q9Z084	Q9z084 tobacco lea
12	271	74.5	190	12	Q9W827	Q9w827 tobacco lea
13	271	74.5	208	12	Q9Z0C4	Q9z0c4 tobacco lea
14	269	73.9	208	12	Q9Z0C0	Q9z0c0 tobacco lea
15	269	73.9	208	12	Q9Z0B8	Q9z0b8 tobacco lea
16	267	73.4	208	12	Q9Z0B6	Q9z0b6 tobacco lea
17	266	73.1	203	12	Q9Z083	Q9z083 tobacco lea
18	266	73.1	363	12	O73577	O73577 cotton leaf
19	259	71.2	208	12	Q9Z0A0	Q9z0a0 tobacco lea
20	259	71.2	363	12	O72705	O72705 cotton leaf
21	259	71.2	363	12	O72719	O72719 cotton leaf
22	258	70.9	190	12	Q9Z086	Q9z086 tobacco lea
23	258	70.9	208	12	Q9Z0C6	Q9z0c6 tobacco lea
24	257	70.6	349	12	Q88888	Q88888 tomato pseu
25	257	70.6	362	12	O56816	O56816 chayote mos
26	256	70.3	359	12	Q9YZV4	Q9yzv4 tomato yell
27	256	70.3	359	12	Q9YZV2	Q9yzv2 tomato yell
28	256	70.3	359	12	Q9YUX7	Q9yux7 tomato yell
29	256	70.3	359	12	Q9YL27	Q9yl27 tomato yell
30	255	70.1	353	12	O72692	O72692 beet curly
31	255	70.1	354	12	Q65438	Q65438 beet curly
32	255	70.1	359	12	Q88942	Q88942 tomato yell
33	254	69.8	190	12	Q9Z0A7	Q9z0a7 tobacco lea
34	254	69.8	362	12	Q9IN48	Q9in48 okra enatio
35	253	69.5	307	12	Q9IET7	Q9iet7 cotton leaf
36	253	69.5	361	12	O72723	O72723 cotton leaf
37	252	69.2	231	12	Q96620	Q96620 african tom
38	252	69.2	354	12	Q9IN42	Q9in42 south afric
39	251	69.0	358	12	Q65418	Q65418 beet curly
40	251	69.0	358	12	Q9JEA2	Q9jea2 cassava gem
41	247	67.9	362	12	Q9YYT7	Q9yyt7 althea rose
42	247	67.9	363	12	O73494	O73494 okra yellow
43	245	67.3	306	12	Q9IET1	Q9iet1 cotton leaf
44	245	67.3	358	12	Q9WR17	Q9wr17 african cas
45	245	67.3	358	12	Q9JEA9	Q9jea9 cassava gem

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 81.3%; Score 296; DB 12; Length 226;
Best Local Similarity 81.4%; Pred. No. 1e-24;
Matches 57; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||||||:||| || ||||||||||  |  |||||:||:|||||||||||||||
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226     226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             81.3%;  Score 296;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 1e-24;
  Matches   54;  Conservative   12;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:|||||||||||||||||::||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             80.8%;  Score 294;  DB 12;  Length 361;
  Best Local Similarity   80.6%;  Pred. No. 2.8e-24;
  Matches   54;  Conservative    9;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || |||||||||| ||:||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225     225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             80.2%;  Score 292;  DB 12;  Length 225;
  Best Local Similarity   80.6%;  Pred. No. 2.7e-24;
  Matches   54;  Conservative    8;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| || ||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             80.2%;  Score 292;  DB 12;  Length 364;
  Best Local Similarity   83.8%;  Pred. No. 4.7e-24;
  Matches   57;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||||||||||||||| ||:|||||||||| ||| ||||||||:|||||||||||| |||
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 DRIFDKTP 68
          ||||   |
Db    170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             79.4%;  Score 289;  DB 12;  Length 185;
  Best Local Similarity   74.3%;  Pred. No. 4.7e-24;
  Matches   52;  Conservative   12;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||:|||||||||: :||:||:|||||:||:||||:|||:||:
Db     89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 149;
  Best Local Similarity   72.9%;  Pred. No. 2.7e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 233;
  Best Local Similarity   72.9%;  Pred. No. 4.5e-23;
```

```
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||||||||: :|| |::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   9
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234     234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             75.5%;  Score 275;  DB 12;  Length 234;
  Best Local Similarity   71.4%;  Pred. No. 2e-22;
  Matches   50;  Conservative   13;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||:|||||||||: :||||::||:||:|||:|||:|||:|:|
Db    110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy     61 DRIFDKTPEP 70
          |||| | | |
Db    170 DRIFMKAPNP 179


RESULT  10
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             74.7%;  Score 272;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 3.4e-22;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             74.7%;  Score 272;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 3.4e-22;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             74.5%;  Score 271;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 4.4e-22;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| ||||||:|| ||| |||||:|: ::|||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  13
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             74.5%;  Score 271;  DB 12;  Length 208;
  Best Local Similarity   62.4%;  Pred. No. 4.9e-22;
  Matches   53;  Conservative    8;  Mismatches    8;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| ||||||:|| ||| |||||:|: ::|||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  14
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             73.9%;  Score 269;  DB 12;  Length 208;
  Best Local Similarity   61.2%;  Pred. No. 8.1e-22;
  Matches   52;  Conservative    9;  Mismatches    8;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||| :|| |||||:|: ::|||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;
```

Query Match 73.9%; Score 269; DB 12; Length 208;
Best Local Similarity 61.2%; Pred. No. 8.1e-22;
Matches 52; Conservative 9; Mismatches 8; Indels 16; Gaps 1;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| ||||||||| :|| |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186
```

Search completed: February 3, 2001, 02:22:50
Job time: 1887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:13:48 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	364	100.0	352	1	VAL1_TGMV	P03567 tomato gold
2	264	72.5	361	1	VAL1_PYMVV	P27258 potato yell
3	251	69.0	358	1	VAL1_CLVK	P14982 cassava lat
4	251	69.0	358	1	VAL1_CLVN	P14972 cassava lat
5	244	67.0	362	1	VAL1_TYLCA	P36279 tomato yell
6	237	65.1	359	1	VAL1_TYLCU	P38609 tomato yell
7	233	64.0	349	1	VAL1_PHUV	Q06923 pepper huas
8	230	63.2	359	1	VAL1_TYLCM	P27260 tomato yell
9	228	62.6	353	1	VAL1_BGMV	P05175 bean golden
10	227	62.4	355	1	VAL1_ABMVW	P21947 abutilon mo
11	224	61.5	358	1	VAL1_BCTV	P14991 beet curly
12	224	61.5	361	1	VAL1_TMOV	Q06657 tomato mott
13	215	59.1	357	1	VAL1_TYLCV	P27259 tomato yell
14	124	34.1	347	1	VAL1_SLCV	P29048 squash leaf
15	64.5	17.7	299	1	Y175_HELPJ	Q9zmq7 helicobacte
16	64.5	17.7	1713	1	LMA3_HUMAN	Q16787 homo sapien
17	63.5	17.4	1610	1	CCAD_MESAU	Q99244 mesocricetu
18	63.5	17.4	2161	1	CCAD_HUMAN	Q01668 homo sapien
19	63.5	17.4	2203	1	CCAD_RAT	P27732 rattus norv
20	62	17.0	295	1	VAL1_TYDVA	P31617 tobacco yel
21	61	16.8	1483	1	CYP1_YEAST	P12351 saccharomyc
22	60	16.5	447	1	CDSA_DROME	P56079 d phosphati
23	59	16.2	387	1	Y4PF_RHISN	P55615 rhizobium s
24	58.5	16.1	136	1	Y452_CAEEL	O62250 caenorhabdi
25	58.5	16.1	297	1	RRPP_RABVP	P06747 rabies viru
26	58.5	16.1	299	1	Y175_HELPY	P56112 helicobacte
27	58.5	16.1	428	1	GBA1_CANAL	P28868 candida alb
28	58.5	16.1	589	1	YE67_HAEIN	P45221 haemophilus
29	58	15.9	617	1	YACH_ECOLI	P36682 escherichia
30	58	15.9	874	1	SLAP_BACLI	P49052 bacillus li
31	57.5	15.8	334	1	G3P_BACST	P00362 bacillus st
32	57.5	15.8	335	1	G3P_BACCO	P15115 bacillus co
33	57.5	15.8	367	1	LHX4_MOUSE	P53776 mus musculu
34	57.5	15.8	419	1	Y223_MYCPN	P75465 mycoplasma
35	57.5	15.8	1852	1	CCAS_CYPCA	P22316 cyprinus ca
36	57	15.7	207	1	IL6_MARMO	O35736 marmota mon
37	57	15.7	256	1	YKJ9_YEAST	P34247 saccharomyc
38	57	15.7	355	1	CRTB_RHOSH	P54905 rhodobacter
39	56.5	15.5	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
40	56.5	15.5	2190	1	CCAD_CHICK	O73700 gallus gall
41	56	15.4	336	1	Y625_METJA	Q58042 methanococc
42	56	15.4	387	1	VE2_HPV41	P27552 human papil
43	56	15.4	513	1	HEM0_CHICK	P18080 gallus gall
44	56	15.4	703	1	UGS2_RAT	P17625 rattus norv
45	56	15.4	836	1	GCSR_HUMAN	Q99062 homo sapien

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 100.0%; Score 364; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.2e-33;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             72.5%;  Score 264;  DB 1;  Length 361;
  Best Local Similarity   71.0%;  Pred. No. 2e-22;
  Matches   49;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             69.0%;  Score 251;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 5.8e-21;
  Matches   45;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             69.0%;  Score 251;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 5.8e-21;
  Matches   45;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             67.0%;  Score 244;  DB 1;  Length 362;
  Best Local Similarity   56.5%;  Pred. No. 3.6e-20;
  Matches   48;  Conservative    9;  Mismatches   12;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |: ||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI---------------FDKTPE 69
           |||               ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             65.1%;  Score 237;  DB 1;  Length 359;
  Best Local Similarity   62.3%;  Pred. No. 2.2e-19;
  Matches   43;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             64.0%;  Score 233;  DB 1;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 5.9e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| |::|| |:|||::| |||||||:|: |: :  |||| ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             63.2%;  Score 230;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 1.3e-18;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV     STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             62.6%;  Score 228;  DB 1;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 2.2e-18;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:|||||||||| |::|| |:|||| | | || |::|: |: |:| || ||: |||
Db    110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             62.4%;  Score 227;  DB 1;  Length 355;
  Best Local Similarity   60.0%;  Pred. No. 2.9e-18;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: |:|||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             61.5%;  Score 224;  DB 1;  Length 358;
  Best Local Similarity   57.1%;  Pred. No. 6.3e-18;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:|||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             61.5%;  Score 224;  DB 1;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 6.3e-18;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             59.1%;  Score 215;  DB 1;  Length 357;
  Best Local Similarity   67.2%;  Pred. No. 6.4e-17;
  Matches   41;  Conservative   10;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |:|||||||| |::||| ||||| || ||| |::|| |: |: ||||||||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.1%;  Score 124;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 1e-06;
  Matches   25;  Conservative   14;  Mismatches   23;  Indels    4;  Gaps    1;

Qy      5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
          |:::| |    |     :|    |:|| |  ||| ||:   |: ::  :||| :|:::|:|
Db    116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy     65 DKTPEP 70
            | |||
Db    172 QKPPEP 177


RESULT  15
Y175_HELPJ
ID   Y175_HELPJ     STANDARD;      PRT;   299 AA.
AC   Q9ZMQ7;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN HP0175 PRECURSOR.
GN   HP0175 OR JHP0161.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC       STRONG, TO C.JEJUNI CBF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001454; AAD05744.1; -.
DR   HSSP; Q13526; 1PIN.
DR   INTERPRO; IPR000297; -.
DR   PFAM; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
KW   Hypothetical protein; Isomerase; Rotamase; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    299       HYPOTHETICAL PROTEIN HP0175.
FT   DOMAIN     154    253       PPIC-LIKE.
SQ   SEQUENCE   299 AA;  34040 MW;  9C037B1CD1110143 CRC64;


  Query Match             17.7%;  Score 64.5;  DB 1;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 4.1;
  Matches   18;  Conservative   14;  Mismatches   11;  Indels   21;  Gaps    2;
```

```
Qy      22 NDAAAEALNAS------------------SKEEALQIIREKIPEKYLFQFHNLNSNLDR 62
           |:| || || :                   :|::| :: : :|||| :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154
```

Search completed: February 3, 2001, 02:24:24
Job time: 636 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 01:40:48 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	364	100.0	352	1	QQCVL1	AL1 protein - toma
2	264	72.5	361	1	QQCVPT	AL1 protein - pota
3	251	69.0	358	2	S07594	hypothetical prote
4	244	67.0	362	1	JQ1887	AL1 protein - toma
5	237	65.1	359	2	S39211	gene C1 protein -
6	233	64.0	349	2	JQ2300	replicase - pepper
7	233	64.0	349	2	S31875	AL1 protein - pepp
8	230	63.2	359	2	S22593	hypothetical prote
9	227	62.4	351	2	JQ2327	AL1 protein - Indi
10	227	62.4	355	1	QQCVW1	AV1 protein - abut
11	224	61.5	358	1	JQ1870	AL1 protein - toma
12	224	61.5	359	2	S39235	gene C1 protein -
13	224	61.5	385	2	S28360	AL1 protein - beet
14	217	59.6	360	2	S59885	replication-associ
15	215	59.1	357	1	QQCVC1	AL1 protein - toma
16	142	39.0	131	2	S45059	AC1 protein (clone
17	124	34.1	347	1	QQCVS1	AL1 protein - squa
18	68	18.7	587	2	JC1419	Fc gamma (IgG) rec
19	64.5	17.7	299	2	B71967	probable peptidyl-
20	64.5	17.7	1713	2	A55347	adhesive ligand ep
21	63.5	17.4	1610	2	A46227	voltage-dependent
22	63.5	17.4	1646	2	JH0422	voltage-dependent
23	63.5	17.4	2161	2	JH0564	calcium channel al
24	63.5	17.4	2181	2	A38198	calcium channel al
25	63.5	17.4	2203	2	T42742	voltage-dependent
26	62.5	17.2	447	2	T12544	hypothetical prote
27	62	17.0	295	2	D42452	C1 protein - tobac
28	61.5	16.9	481	2	A70091	probable phosphoes
29	61	16.8	840	2	T36175	probable large ATP
30	61	16.8	1502	1	RGBYH1	CYC1/CYP3 transcri
31	60.5	16.6	354	2	A75087	acetyl ornithine d
32	60.5	16.6	1229	2	T48959	kinesin-like prote
33	60	16.5	447	2	S52437	CDP-diacylglycerol
34	59.5	16.3	2137	2	T05244	hypothetical prote
35	59	16.2	160	2	G82060	hypothetical prote
36	59	16.2	316	2	C82085	conserved hypothet
37	59	16.2	397	2	B71078	probable NADH oxid
38	59	16.2	416	2	A82892	hypothetical prote
39	58.5	16.1	136	2	T22240	hypothetical prote
40	58.5	16.1	297	1	MNVNRV	nonstructural prot
41	58.5	16.1	299	2	G64541	cell binding facto
42	58.5	16.1	589	2	D64125	hypothetical prote
43	58.5	16.1	1044	2	T43800	protein kinase bub
44	58.5	16.1	1265	2	T47626	structural mainten
45	58.5	16.1	1743	2	T15893	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 100.0%; Score 364; DB 1; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.4e-33;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match            72.5%;  Score 264;  DB 1;  Length 361;
  Best Local Similarity  71.0%;  Pred. No. 4.2e-22;
  Matches   49;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:||||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match            69.0%;  Score 251;  DB 2;  Length 358;
  Best Local Similarity  64.3%;  Pred. No. 1.2e-20;
  Matches   45;  Conservative   13;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match            67.0%;  Score 244;  DB 1;  Length 362;
  Best Local Similarity  56.5%;  Pred. No. 7.5e-20;
  Matches   48;  Conservative    9;  Mismatches   12;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |: ||||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match            65.1%;  Score 237;  DB 2;  Length 359;
  Best Local Similarity  62.3%;  Pred. No. 4.5e-19;
  Matches   43;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match            64.0%;  Score 233;  DB 2;  Length 349;
  Best Local Similarity  61.4%;  Pred. No. 1.2e-18;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.0%;  Score 233;  DB 2;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 1.2e-18;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::|| |:|||::| ||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.2%;  Score 230;  DB 2;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 2.8e-18;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.4%;  Score 227;  DB 2;  Length 351;
  Best Local Similarity   64.2%;  Pred. No. 5.9e-18;
  Matches   43;  Conservative    9;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| | |||: || |||:|:||  |  ||  ||:::|||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.4%;  Score 227;  DB 1;  Length 355;
  Best Local Similarity   60.0%;  Pred. No. 6e-18;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: |:||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 224;  DB 1;  Length 358;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-17;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;
```

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176
```

RESULT 12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             61.5%;  Score 224;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.3e-17;
  Matches   40;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|:: |  |:  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179
```

RESULT 13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             61.5%;  Score 224;  DB 2;  Length 385;
  Best Local Similarity   57.1%;  Pred. No. 1.4e-17;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| ||:||: |:|||||:| ::|||||::|: |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
           :|| : |:|
Db    197 QKIFQRPPDP 206
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.6%;  Score 217;  DB 2;  Length 360;
  Best Local Similarity   62.1%;  Pred. No. 8.1e-17;
  Matches   41;  Conservative   12;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| ||:|||||||| |::||| |||:|: ||  || |:||| |: :: ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.1%;  Score 215;  DB 1;  Length 357;
  Best Local Similarity   67.2%;  Pred. No. 1.3e-16;
  Matches   41;  Conservative   10;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |:|||||||| |::||| |||||: || ||| |:::| |: |: ||||||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171
```

Search completed: February 3, 2001, 02:17:36
Job time: 2208 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 01:51:18 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	230	63.2	359	3	US-08-809-103B-2	Sequence 2, Appli
2	230	63.2	359	3	US-08-809-103B-4	Sequence 4, Appli
3	230	63.2	359	3	US-08-809-103B-6	Sequence 6, Appli
4	230	63.2	359	3	US-08-809-103B-8	Sequence 8, Appli
5	64.5	17.7	1713	3	US-08-600-982-24	Sequence 24, Appl
6	64.5	17.7	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
7	63.5	17.4	2161	1	US-07-745-206A-2	Sequence 2, Appli
8	63.5	17.4	2161	1	US-08-455-543A-49	Sequence 49, Appl
9	63.5	17.4	2161	1	US-08-455-543A-51	Sequence 51, Appl
10	63.5	17.4	2161	2	US-08-223-305C-49	Sequence 49, Appl
11	63.5	17.4	2161	2	US-08-223-305C-51	Sequence 51, Appl
12	63.5	17.4	2161	2	US-08-311-363-2	Sequence 2, Appli
13	60	16.5	446	2	US-08-672-814D-11	Sequence 11, Appl
14	58	15.9	844	3	US-09-029-267-20	Sequence 20, Appl
15	57.5	15.8	439	1	US-07-637-870-9	Sequence 9, Appli
16	57.5	15.8	439	1	US-07-637-399-6	Sequence 6, Appli
17	57.5	15.8	439	1	US-08-112-703-6	Sequence 6, Appli
18	56.5	15.5	334	5	5290690-11	Patent No. 5290690
19	56	15.4	454	3	US-08-446-100-31	Sequence 31, Appl
20	56	15.4	602	2	US-08-419-652-6	Sequence 6, Appli
21	56	15.4	771	1	US-07-923-976-6	Sequence 6, Appli
22	56	15.4	783	5	5422248-2	Patent No. 5422248
23	56	15.4	836	1	US-07-923-976-4	Sequence 4, Appli
24	56	15.4	863	1	US-07-923-976-8	Sequence 8, Appli
25	54	14.8	454	3	US-08-446-100-30	Sequence 30, Appl
26	53	14.6	266	2	US-07-857-224B-40	Sequence 40, Appl
27	53	14.6	454	3	US-08-446-100-26	Sequence 26, Appl
28	53	14.6	454	3	US-08-446-100-27	Sequence 27, Appl
29	53	14.6	454	3	US-08-446-100-28	Sequence 28, Appl
30	53	14.6	454	3	US-08-446-100-29	Sequence 29, Appl
31	53	14.6	589	3	US-08-942-008-2	Sequence 2, Appli
32	53	14.6	783	5	5231168-2	Patent No. 5231168
33	52.5	14.4	81	3	US-08-995-159-6	Sequence 6, Appli
34	52.5	14.4	303	3	US-09-045-973-1	Sequence 1, Appli
35	52.5	14.4	312	1	US-08-414-625-2	Sequence 2, Appli
36	52	14.3	439	1	US-07-762-681A-1	Sequence 1, Appli
37	52	14.3	482	1	US-07-792-885A-1	Sequence 1, Appli
38	52	14.3	482	1	US-08-142-439A-7	Sequence 7, Appli
39	52	14.3	482	2	US-08-869-477-7	Sequence 7, Appli
40	52	14.3	751	3	US-09-036-987A-24	Sequence 24, Appl
41	51.5	14.1	373	3	US-08-258-287B-42	Sequence 42, Appl
42	51.5	14.1	373	3	US-08-368-704C-42	Sequence 42, Appl
43	51.5	14.1	834	1	US-07-977-434-8	Sequence 8, Appli
44	51.5	14.1	834	1	US-08-458-819-8	Sequence 8, Appli
45	51.5	14.1	834	4	PCT-US91-07035-8	Sequence 8, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match           63.2%;  Score 230;  DB 3;  Length 359;
  Best Local Similarity  59.4%;  Pred. No. 1.3e-22;
  Matches  41;  Conservative  11;  Mismatches  17;  Indels  0;  Gaps  0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             63.2%;  Score 230;  DB 3;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 1.3e-22;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             63.2%;  Score 230;  DB 3;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 1.3e-22;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             63.2%;  Score 230;  DB 3;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 1.3e-22;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             17.7%;  Score 64.5;  DB 3;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 9.2;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   6
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24
```

```
  Query Match             17.7%;  Score 64.5;  DB 4;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 9.2;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   7
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             17.4%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 17;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   8
US-08-455-543A-49
;
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             17.4%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 17;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;
```

```
Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  9
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match             17.4%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 17;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  10
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
```

```
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             17.4%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 17;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |    ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  11
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             17.4%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 17;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |    ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  12
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,363
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-51506
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (619)238-0999
;       TELEFAX:  (619)238-0062
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2161 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             17.4%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 17;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  13
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
;  GENERAL INFORMATION:
;    APPLICANT:  Leung, David W.
;    APPLICANT:  Weeks, Reitha
;    TITLE OF INVENTION:  MAMMALIAN
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cell Therapeutics, Inc.
;      STREET:  201 Elliott Avenue West
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98119
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
;      COMPUTER:  PC Clone (486 microprocessor)
;      OPERATING SYSTEM:  MS-DOS Version 6.1, Windows NT
;      SOFTWARE:  WORD 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,814D
;      FILING DATE:  28-Jun-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Faciszewski, Stephen
;      REGISTRATION NUMBER:  36,131
;      REFERENCE/DOCKET NUMBER:  1803
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)282-7100
;      TELEFAX:  (206)284-6206
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  446
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:  CDP-diacylglycerol synthase (CDS)
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-672-814D-11


  Query Match             16.5%;  Score 60;  DB 2;  Length 446;
  Best Local Similarity   34.0%;  Pred. No. 6.2;
  Matches   17;  Conservative    9;  Mismatches   22;  Indels    2;  Gaps    1;

Qy      20 TSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           : :||| :  :|:   : :    |||||:       |  || : |||||
Db      20 SGSDAANKRNSAADSSDHVDSEEEKIPEEKFVD--ELAKNLPQGTDKTPE 67


RESULT  14
US-09-029-267-20
; Sequence 20, Application US/09029267
; Patent No. 6107057
;  GENERAL INFORMATION:
;    APPLICANT:  Crawford, Kenneth
;    APPLICANT:  Zaror, Isabel
;    APPLICANT:  Innis, Michael
;    TITLE OF INVENTION:  Pichia Secretory Leader for Protein
;    TITLE OF INVENTION:  Expression
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/029,267
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Chung, Ling-Fong
;      REGISTRATION NUMBER:  36,482
;      REFERENCE/DOCKET NUMBER:  1165.100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 601-2704
;      TELEFAX:  (510) 655-3542
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  844 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-029-267-20


  Query Match             15.9%;  Score 58;  DB 3;  Length 844;
  Best Local Similarity   27.5%;  Pred. No. 26;
  Matches   11;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      24 AAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           |  :  |: :: |:||::  :  ||  : |  | :| :|::
Db     714 AHSTTLSVATYEQALEMSNQYAPEHLILQIENASSYVDQV 753


RESULT  15
```

```
US-07-637-870-9
; Sequence 9, Application US/07637870
; Patent No. 5310665
;  GENERAL INFORMATION:
;    APPLICANT:  Lambier, Anne-Marie
;    APPLICANT:  Lasters, Ignace
;    APPLICANT:  Quax, Wilhemus J.
;    APPLICANT:  Van Der Laan, Jan M.
;    TITLE OF INVENTION:  NOVEL GLUCOSE ISOMERASES HAVING
;    TITLE OF INVENTION:  ALTERED SUBSTRATE SPECIFICITY
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morrison & Foerster
;      STREET:  545 Middlefield Road, Suite 200
;      CITY:  Menlo Park
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/637,870
;      FILING DATE:  19910104
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H.
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  24615-20019.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-813-5600
;      TELEFAX:  415-327-2951
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  439 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-637-870-9


  Query Match             15.8%;  Score 57.5;  DB 1;  Length 439;
  Best Local Similarity   25.6%;  Pred. No. 13;
  Matches   20;  Conservative   11;  Mismatches   30;  Indels   17;  Gaps    3;

Qy       4 WGEFQVDGRSARGGCQTSN----DAAAEALNASSKEEALQIIREKIPEKYLFQFHNLN-- 57
           |    :||    ||         :  :|   |  : :|   | ||:  || : ||:::
Db      48 WHTITMDGSDPFGGATMERPWDLEGGSELDRAHRRVDAFFEIAEKLGVKY-YCFHDIDIA 106

Qy      58 ----------SNLDRIFD 65
                    :||| | |
Db     107 PTGNSLKEFYANLDEITD 124


Search completed: February  3, 2001, 02:19:38
Job time: 1700 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 01:18:13 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-1
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	230	63.2	353	18	W34338	Bean golden mosaic
2	230	63.2	353	18	W34332	Bean golden mosaic
3	230	63.2	353	18	W34333	Bean golden mosaic
4	230	63.2	353	18	W34334	Bean golden mosaic
5	230	63.2	353	18	W34335	Bean golden mosaic
6	230	63.2	359	17	R88870	Sardinian tomato y
7	230	63.2	359	17	R88871	Sardinian tomato y
8	230	63.2	359	17	R88872	Sardinian tomato y
9	228	62.6	353	8	P70407	ORF 4 gene product
10	228	62.6	361	18	W34336	Tomato mottle viru
11	228	62.6	361	18	W34324	Tomato mottle viru
12	228	62.6	361	18	W34325	Tomato mottle viru
13	228	62.6	361	18	W34326	Tomato mottle viru
14	221.5	60.9	361	8	P70562	Product of ORF 4 f
15	216	59.3	362	19	W56495	Tobacco leaf curli
16	215	59.1	357	18	W34329	Tomato yellow leaf
17	215	59.1	357	18	W34330	Tomato yellow leaf
18	215	59.1	357	18	W34331	Tomato yellow leaf
19	207	56.9	357	18	W34337	Tomato yellow leaf
20	67.5	18.5	512	19	W68473	HIV-1 strain YBF30
21	65	17.9	292	21	Y92317	Geminivirus Rep C1
22	65	17.9	335	21	Y92318	Mastrevirus Rep in
23	64.5	17.7	299	18	W55450	H. pylori ORF 02ae
24	64.5	17.7	1713	16	R70148	Deduced sequence o
25	63.5	17.4	619	13	R27651	Human calcium chan
26	63.5	17.4	2161	14	R33545	Sequence of the al
27	63.5	17.4	2161	16	R71001	Human neuronal cal
28	63.5	17.4	2161	16	R71002	Human neuronal cal
29	63.5	17.4	2161	19	W63137	Human calcium chan
30	63.5	17.4	2161	19	W63149	Human calcium chan
31	62.5	17.2	447	21	Y87089	Human secreted pro
32	62.5	17.2	447	21	Y66744	Membrane-bound pro
33	62.5	17.2	447	21	Y50947	Human adult aorta
34	62.5	17.2	456	21	Y87190	Human secreted pro
35	60.5	16.6	131	18	W34327	Tomato mottle viru
36	60.5	16.6	204	20	W89814	Protein encoded by
37	60.5	16.6	299	20	W89829	Protein encoded by
38	60.5	16.6	299	20	W89849	Antigen from clust
39	58.5	16.1	297	8	P70318	Protein M1 encoded
40	58.5	16.1	299	19	W98322	H. pylori GHP0 136
41	58.5	16.1	299	19	W73034	Helicobacter pylor
42	58	15.9	318	20	Y31895	Soybean cyclin del
43	58	15.9	844	18	W09773	Histidinol dehydro
44	57.5	15.8	188	21	Y95043	Candida albicans p
45	57	15.7	108	16	R82847	Murine mutant cyto

ALIGNMENTS

RESULT 1
W34338
ID W34338 standard; Protein; 353 AA.
XX
AC W34338;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 protein.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93314.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

```
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             63.2%;  Score 230;  DB 18;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   2
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             63.2%;  Score 230;  DB 18;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   3
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
```

```
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            63.2%;  Score 230;  DB 18;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179


RESULT   4
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            63.2%;  Score 230;  DB 18;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179


RESULT   5
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            63.2%;  Score 230;  DB 18;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 6.5e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169
```

```
Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             63.2%;  Score 230;  DB 17;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 6.6e-21;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   7
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             63.2%;  Score 230;  DB 17;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 6.6e-21;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   8
```

```
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             63.2%;  Score 230;  DB 17;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 6.6e-21;
  Matches   41;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   9
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             62.6%;  Score 228;  DB 8;  Length 353;
  Best Local Similarity   61.4%;  Pred. No. 1.2e-20;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             62.6%;  Score 228;  DB 18;  Length 361;
  Best Local Similarity   58.6%;  Pred. No. 1.2e-20;
  Matches   41;  Conservative   16;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             62.6%;  Score 228;  DB 18;  Length 361;
  Best Local Similarity   58.6%;  Pred. No. 1.2e-20;
  Matches   41;  Conservative   16;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             62.6%;  Score 228;  DB 18;  Length 361;
  Best Local Similarity   58.6%;  Pred. No. 1.2e-20;
  Matches   41;  Conservative   16;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             62.6%;  Score 228;  DB 18;  Length 361;
  Best Local Similarity   58.6%;  Pred. No. 1.2e-20;
  Matches   41;  Conservative   16;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             60.9%;  Score 221.5;  DB 8;  Length 361;
  Best Local Similarity   67.7%;  Pred. No. 7.8e-20;
  Matches   44;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           | |||||||||||  |::||| |||||: || :|| |:||| |::|: |||||| || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             59.3%;  Score 216;  DB 19;  Length 362;
  Best Local Similarity   53.6%;  Pred. No. 3.8e-19;
  Matches   45;  Conservative   10;  Mismatches   13;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | :| |||||||||  ||::||| |||:|: ||  || |:||| |: :: ||||||||||
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 RI---------------FDKTPE 69
           ||               ||: ||
Db     171 rifapplevfvcpfssssfdqvpe 194



Search completed: February  3, 2001, 02:15:25
Job time: 3432 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:50 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	281	77.6	226	12	O09727	O09727 leonurus mo
2	281	77.6	226	12	Q9WHF6	Q9whf6 tomato mild
3	279	77.1	361	12	Q67574	Q67574 bean golden
4	277	76.5	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	277	76.5	364	12	Q9QS55	Q9qs55 sweet potat
6	274	75.7	185	12	Q98693	Q98693 sida golden
7	266	73.5	149	12	P88975	P88975 macroptiliu
8	266	73.5	233	12	Q9YLA4	Q9yla4 macroptiliu
9	260	71.8	234	12	O39180	O39180 potato yell
10	257	71.0	190	12	Q9Z089	Q9z089 tobacco lea
11	257	71.0	190	12	Q9Z084	Q9z084 tobacco lea
12	256	70.7	190	12	Q9W827	Q9w827 tobacco lea
13	256	70.7	208	12	Q9Z0C4	Q9z0c4 tobacco lea
14	254	70.2	208	12	Q9Z0C0	Q9z0c0 tobacco lea
15	254	70.2	208	12	Q9Z0B8	Q9z0b8 tobacco lea
16	252	69.6	208	12	Q9Z0B6	Q9z0b6 tobacco lea
17	251	69.3	203	12	Q9Z083	Q9z083 tobacco lea
18	251	69.3	363	12	O73577	O73577 cotton leaf
19	244	67.4	208	12	Q9Z0A0	Q9z0a0 tobacco lea
20	244	67.4	363	12	O72705	O72705 cotton leaf
21	244	67.4	363	12	O72719	O72719 cotton leaf
22	243	67.1	190	12	Q9Z086	Q9z086 tobacco lea
23	243	67.1	208	12	Q9Z0C6	Q9z0c6 tobacco lea
24	242	66.9	349	12	Q88888	Q88888 tomato pseu
25	242	66.9	362	12	O56816	O56816 chayote mos
26	241	66.6	359	12	Q9YZV4	Q9yzv4 tomato yell
27	241	66.6	359	12	Q9YZV2	Q9yzv2 tomato yell
28	241	66.6	359	12	Q9YUX7	Q9yux7 tomato yell
29	241	66.6	359	12	Q9YL27	Q9yl27 tomato yell
30	240	66.3	353	12	O72692	O72692 beet curly
31	240	66.3	354	12	Q65438	Q65438 beet curly
32	240	66.3	359	12	Q88942	Q88942 tomato yell
33	239	66.0	190	12	Q9Z0A7	Q9z0a7 tobacco lea
34	239	66.0	362	12	Q9IN48	Q9in48 okra enatio
35	238	65.7	307	12	Q9IET7	Q9iet7 cotton leaf
36	238	65.7	361	12	O72723	O72723 cotton leaf
37	237	65.5	231	12	Q96620	Q96620 african tom
38	237	65.5	354	12	Q9IN42	Q9in42 south afric
39	236	65.2	358	12	Q65418	Q65418 beet curly
40	236	65.2	358	12	Q9JEA2	Q9jea2 cassava gem
41	232	64.1	362	12	Q9YYT7	Q9yyt7 althea rose
42	232	64.1	363	12	O73494	O73494 okra yellow
43	230	63.5	306	12	Q9IET1	Q9iet1 cotton leaf
44	230	63.5	358	12	Q9WR17	Q9wr17 african cas
45	230	63.5	358	12	Q9JEA9	Q9jea9 cassava gem

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.6%; Score 281; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 6.7e-23;
Matches 54; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||||| :: || || |||||||||| | |||||:||:||||||||||||||||
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             77.6%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 6.7e-23;
  Matches   51;  Conservative   13;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|| :| || ||:||||||||||||||||::||:||:|:::|||:|||:|||
Db    111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.1%;  Score 279;  DB 12;  Length 361;
  Best Local Similarity   76.1%;  Pred. No. 1.9e-22;
  Matches   51;  Conservative   10;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          || ||||| :| || ||:||||:||||||||||||:|||:||:|||:|||:||||:||||||
Db    113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |:|
Db    173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.5%;  Score 277;  DB 12;  Length 225;
  Best Local Similarity   76.1%;  Pred. No. 1.8e-22;
  Matches   51;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          || ||:|| :| || || ||||:||||||||||||:|||:||:|||:|||:||||:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             76.5%;  Score 277;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 3.1e-22;
  Matches   54;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||| :| || ||:|||||||||| ||| ||||||||:|||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.7%;  Score 274;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 3e-22;
  Matches   49;  Conservative   13;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || ||:|||||||||: :||:||:|||||:||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.5%;  Score 266;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 1.7e-21;
  Matches   48;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || ||||||||||||: :|| |::|:::||:|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.5%;  Score 266;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 2.9e-21;
```

```
  Matches   48;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|| :| || |||||||||||: :|| |::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   9
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             71.8%;  Score 260;  DB 12;  Length 234;
  Best Local Similarity   67.1%;  Pred. No. 1.3e-20;
  Matches   47;  Conservative   14;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|| :| || ||:|||||||||: :||||::||:||:|||:|||:|||:|:|
Db    110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy     61 DRIFDKTPEP 70
          |||| | | |
Db    170 DRIFMKAPNP 179


RESULT  10
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             71.0%;  Score 257;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.2e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|| :| |||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             71.0%;  Score 257;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.2e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|| :| |||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             70.7%;  Score 256;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.8e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||||| :| |||| :||| ||||||:|| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  13
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             70.7%;  Score 256;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 3.1e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||||| :| |||| :||| ||||||:|| ||| |||||:|: ::||:|||||||
Db    102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  14
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             70.2%;  Score 254;  DB 12;  Length 208;
  Best Local Similarity   57.6%;  Pred. No. 5.1e-20;
  Matches   49;  Conservative   10;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|| :| |||| :||| ||||||||| :|| |||||:|: ::||:|||||||
Db    102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;
```

Query Match 70.2%; Score 254; DB 12; Length 208;
Best Local Similarity 57.6%; Pred. No. 5.1e-20;
Matches 49; Conservative 10; Mismatches 10; Indels 16; Gaps 1;

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|| :| |||| :||| ||||||||| :|| |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186
```

Search completed: February 3, 2001, 02:22:50
Job time: 1887 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:24 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	96.4	352	1	VAL1_TGMV	P03567 tomato gold
2	249	68.8	361	1	VAL1_PYMVV	P27258 potato yell
3	236	65.2	358	1	VAL1_CLVK	P14982 cassava lat
4	236	65.2	358	1	VAL1_CLVN	P14972 cassava lat
5	229	63.3	362	1	VAL1_TYLCA	P36279 tomato yell
6	222	61.3	359	1	VAL1_TYLCU	P38609 tomato yell
7	218	60.2	349	1	VAL1_PHUV	Q06923 pepper huas
8	215	59.4	359	1	VAL1_TYLCM	P27260 tomato yell
9	213	58.8	353	1	VAL1_BGMV	P05175 bean golden
10	212	58.6	355	1	VAL1_ABMVW	P21947 abutilon mo
11	209	57.7	358	1	VAL1_BCTV	P14991 beet curly
12	209	57.7	361	1	VAL1_TMOV	Q06657 tomato mott
13	200	55.2	357	1	VAL1_TYLCV	P27259 tomato yell
14	124	34.3	347	1	VAL1_SLCV	P29048 squash leaf
15	68.5	18.9	1713	1	LMA3_HUMAN	Q16787 homo sapien
16	65.5	18.1	1610	1	CCAD_MESAU	Q99244 mesocricetu
17	65.5	18.1	2161	1	CCAD_HUMAN	Q01668 homo sapien
18	65.5	18.1	2203	1	CCAD_RAT	P27732 rattus norv
19	64.5	17.8	299	1	Y175_HELPJ	Q9zmq7 helicobacte
20	64.5	17.8	334	1	G3P_BACST	P00362 bacillus st
21	64.5	17.8	335	1	G3P_BACCO	P15115 bacillus co
22	64.5	17.8	557	1	SASB_ANAPL	Q04791 anas platyr
23	62.5	17.3	419	1	Y223_MYCPN	P75465 mycoplasma
24	62	17.1	295	1	VAL1_TYDVA	P31617 tobacco yel
25	62	17.1	1483	1	CYP1_YEAST	P12351 saccharomyc
26	60.5	16.7	136	1	Y452_CAEEL	O62250 caenorhabdi
27	60	16.6	387	1	VE2_HPV41	P27552 human papil
28	60	16.6	447	1	CDSA_DROME	P56079 d phosphati
29	60	16.6	513	1	HEM0_CHICK	P18080 gallus gall
30	59	16.3	129	1	RK12_PORPU	P51339 porphyra pu
31	59	16.3	247	1	YCP4_YEAST	P25349 saccharomyc
32	58.5	16.2	299	1	Y175_HELPY	P56112 helicobacte
33	58.5	16.2	589	1	YE67_HAEIN	P45221 haemophilus
34	58.5	16.2	2190	1	CCAD_CHICK	O73700 gallus gall
35	58	16.0	387	1	Y4PF_RHISN	P55615 rhizobium s
36	58	16.0	617	1	YACH_ECOLI	P36682 escherichia
37	58	16.0	874	1	SLAP_BACLI	P49052 bacillus li
38	57.5	15.9	266	1	ETFB_MYCLE	O33095 mycobacteri
39	57.5	15.9	266	1	ETFB_MYCTU	O53276 mycobacteri
40	57.5	15.9	525	1	P2BB_RAT	P20651 rattus norv
41	57.5	15.9	1852	1	CCAS_CYPCA	P22316 cyprinus ca
42	57	15.7	312	1	SOX3_CHICK	P48433 gallus gall
43	57	15.7	355	1	CRTB_RHOSH	P54905 rhodobacter
44	57	15.7	382	1	IPAC_SHIDY	Q03946 shigella dy
45	57	15.7	393	1	METK_ASCIM	P50304 ascobolus i

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 96.4%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 8.8e-31;
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||| :| ||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             68.8%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 6.6e-20;
  Matches   46;  Conservative   12;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|| :| || || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPEP 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             65.2%;  Score 236;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.7e-18;
  Matches   42;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             65.2%;  Score 236;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.7e-18;
  Matches   42;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             63.3%;  Score 229;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 9.8e-18;
  Matches   45;  Conservative   10;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|| :| || |::||| |:|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             61.3%;  Score 222;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 5.6e-17;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.2%;  Score 218;  DB 1;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 1.5e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|| :| || |::|| |:|||::| ||||||:|: |: :  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             59.4%;  Score 215;  DB 1;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 3.2e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV     STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             58.8%;  Score 213;  DB 1;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 5.2e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             58.6%;  Score 212;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 6.8e-16;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|| :| || ||:||: |:|||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             57.7%;  Score 209;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 1.4e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||:|| :| || ||:||: |:||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             57.7%;  Score 209;  DB 1;  Length 361;
  Best Local Similarity   52.9%;  Pred. No. 1.5e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|| :| || |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :|| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             55.2%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 1.4e-14;
  Matches   38;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:|| :| || |::||| ||||| : || ||| |::|| |: |: ||||| :|||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.3%;  Score 124;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 2.4e-06;
  Matches   25;  Conservative   14;  Mismatches   23;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |: ::  :||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
LMA3_HUMAN
ID   LMA3_HUMAN     STANDARD;      PRT;  1713 AA.
AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   LAMININ ALPHA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KERATINOCYTES;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC       VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC       6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC       PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC       KERATINOCYTES.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC       NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC       AND LAMININ-7 (KS-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC       TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
```

```
CC       A.
CC   -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC       EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC       SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC       AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC       DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC       IN THE CNS.
CC   -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN G IS GLOBULAR.
CC   -!- DISEASE: JUNCTIONAL EPIDERMOLYSIS BULLOSA GRAVIS IS A BLISTERING
CC       DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION OF BASAL
CC       CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED NUMBER OF
CC       HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT MEMBRANE OF
CC       PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS BULLOSA.
CC   -!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   HSSP; P02468; 1TLE.
DR   MIM; 600805; -.
DR   MIM; 226700; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR001791; -.
DR   INTERPRO; IPR002049; -.
DR   PFAM; PF00053; laminin_EGF; 2.
DR   PFAM; PF00054; laminin_G; 2.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     20       POTENTIAL.
FT   CHAIN       21   1713       LAMININ ALPHA-3 CHAIN.
FT   DOMAIN      46    201       DOMAIN III A.
FT   DOMAIN      67    185       2.5 X LAMININ EGF-LIKE REPEATS.
FT   DOMAIN      67    113       LAMININ EGF-LIKE 1.
FT   DOMAIN     114    166       LAMININ EGF-LIKE 2.
FT   DOMAIN     167    185       LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN     202    793       DOMAIN II AND I (HEPTAT REPEATS).
FT   DOMAIN     794   1713       5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT   DOMAIN     794    970       LAMININ G-LIKE 1.
FT   DOMAIN     971   1139       LAMININ G-LIKE 2.
FT   DOMAIN    1140   1353       LAMININ G-LIKE 3.
FT   DOMAIN    1354   1529       LAMININ G-LIKE 4.
FT   DOMAIN    1530   1713       LAMININ G-LIKE 5.
FT   DOMAIN     231    327       COILED COIL (POTENTIAL).
FT   DOMAIN     396    548       COILED COIL (POTENTIAL).
FT   DOMAIN     594    621       COILED COIL (POTENTIAL).
FT   DOMAIN     702    765       COILED COIL (POTENTIAL).
FT   DOMAIN    1686   1713       COILED COIL (POTENTIAL).
FT   DISULFID    67     76       BY SIMILARITY.
FT   DISULFID    69     83       BY SIMILARITY.
FT   DISULFID    86     95       BY SIMILARITY.
FT   DISULFID    98    111       BY SIMILARITY.
FT   DISULFID   114    126       BY SIMILARITY.
FT   DISULFID   116    135       BY SIMILARITY.
FT   DISULFID   137    146       BY SIMILARITY.
FT   DISULFID   149    164       BY SIMILARITY.
FT   DISULFID   202    202       INTERCHAIN (PROBABLE).
FT   DISULFID   205    205       INTERCHAIN (PROBABLE).
FT   SITE       658    660       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD   542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   645    645       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   745    745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   882    882       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   964    964       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1108   1108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1131   1131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1325   1325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1477   1477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1667   1667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1     45       MGWLWIFGAALGQCLGYSSQQQRVPFLQPPGQSQLQASYVE
FT                               FRPS -> KVSSYGGYLTYQAKSFALPGDMVLLEKKPDVQL
FT                               TGQHMSIIYEETNTPRPDRLHHGRVHVVEGNFRHASSRAPV
FT                               SREELMTVLSRLADVRIQGLYFTETQRLTLSEVGLEEASDT
FT                               GSGRIALAVEICACPPAYAGDSC (IN ISOFORM B).
FT   CONFLICT     5      5       W -> R (IN REF. 2).
FT   CONFLICT   123    125       ATG -> GMC (IN REF. 2).
FT   CONFLICT   481    481       M -> K (IN REF. 2).
FT   CONFLICT   754    754       R -> L (IN REF. 2).
FT   CONFLICT   969    969       E -> Q (IN REF. 2).
FT   CONFLICT  1052   1052       D -> A (IN REF. 2).
FT   CONFLICT  1184   1184       G -> A (IN REF. 2).
SQ   SEQUENCE  1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;
```

```
  Query Match             18.9%;  Score 68.5;  DB 1;  Length 1713;
  Best Local Similarity   36.1%;  Pred. No. 15;
  Matches   22;  Conservative   11;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGAAAAGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           || | |       :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


Search completed: February  3, 2001, 02:24:24
Job time: 636 sec
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:36 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	96.4	352	1	QQCVL1	AL1 protein - toma
2	249	68.8	361	1	QQCVPT	AL1 protein - pota
3	236	65.2	358	2	S07594	hypothetical prote
4	229	63.3	362	1	JQ1887	AL1 protein - toma
5	222	61.3	359	2	S39211	gene C1 protein -
6	218	60.2	349	2	JQ2300	replicase - pepper
7	218	60.2	349	2	S31875	AL1 protein - pepp
8	215	59.4	359	2	S22593	hypothetical prote
9	212	58.6	351	2	JQ2327	AL1 protein - Indi
10	212	58.6	355	1	QQCVW1	AV1 protein - abut
11	209	57.7	358	1	JQ1870	AL1 protein - toma
12	209	57.7	359	2	S39235	gene C1 protein -
13	209	57.7	385	2	S28360	AL1 protein - beet
14	202	55.8	360	2	S59885	replication-associ
15	200	55.2	357	1	QQCVC1	AL1 protein - toma
16	142	39.2	131	2	S45059	AC1 protein (clone
17	124	34.3	347	1	QQCVS1	AL1 protein - squa
18	68.5	18.9	1713	2	A55347	adhesive ligand ep
19	65.5	18.1	1610	2	A46227	voltage-dependent
20	65.5	18.1	1646	2	JH0422	voltage-dependent
21	65.5	18.1	2161	2	JH0564	calcium channel al
22	65.5	18.1	2181	2	A38198	calcium channel al
23	65.5	18.1	2203	2	T42742	voltage-dependent
24	65	18.0	371	2	B69502	alcohol dehydrogen
25	64.5	17.8	299	2	B71967	probable peptidyl-
26	64.5	17.8	335	2	JS0164	glyceraldehyde-3-p
27	64.5	17.8	557	2	A47162	thiolesterase B (E
28	63.5	17.5	334	1	DEBSGF	glyceraldehyde-3-p
29	63.5	17.5	392	2	T45290	acetyl-CoA C-acety
30	63	17.4	397	2	B71078	probable NADH oxid
31	62.5	17.3	419	2	S73846	MG223 homolog F10_
32	62	17.1	295	2	D42452	C1 protein - tobac
33	62	17.1	1502	1	RGBYH1	CYC1/CYP3 transcri
34	61.5	17.0	481	2	A70091	probable phosphoes
35	61	16.9	392	2	T44362	acetyl-CoA C-acety
36	61	16.9	587	2	JC1419	Fc gamma (IgG) rec
37	60.5	16.7	136	2	T22240	hypothetical prote
38	60.5	16.7	201	2	A81380	probable flagellar
39	60	16.6	384	1	W2WL41	E2 protein - human
40	60	16.6	447	2	S52437	CDP-diacylglycerol
41	60	16.6	513	1	SYCHLE	5-aminolevulinate
42	59	16.3	129	2	S73260	ribosomal protein
43	59	16.3	247	2	S26733	hypothetical prote
44	58.5	16.2	266	2	T35449	hypothetical prote
45	58.5	16.2	299	2	G64541	cell binding facto

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 96.4%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.1e-30;
Matches 67; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||  :| ||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             68.8%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 1e-19;
  Matches   46;  Conservative   12;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|| :| || || |||||||||: :|| |::||:||:|||:|||:||| || 
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             65.2%;  Score 236;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 2.6e-18;
  Matches   42;  Conservative   14;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|| :| || |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.3%;  Score 229;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 1.5e-17;
  Matches   45;  Conservative   10;  Mismatches   14;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || |||||:|| :| || |::||| |:||| || ||| ::|| |: |: |||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.3%;  Score 222;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 8.9e-17;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|| :| || ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 218;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 2.4e-16;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|| :| || |::||  |:|||::| |||||||:|: |: |:  |||| ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.2%; Score 218; DB 2; Length 349;
Best Local Similarity 57.1%; Pred. No. 2.4e-16;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy 61 DRIFDKTPEP 70
Db 170 NRIFQTPPEP 179

RESULT 8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 59.4%; Score 215; DB 2; Length 359;
Best Local Similarity 55.1%; Pred. No. 5.2e-16;
Matches 38; Conservative 12; Mismatches 19; Indels 0; Gaps 0;

Qy 2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
Db 111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy 62 RIFDKTPEP 70
Db 171 KVFQVPPAP 179

RESULT 9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.6%; Score 212; DB 2; Length 351;
Best Local Similarity 59.7%; Pred. No. 1.1e-15;
Matches 40; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

Qy 4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
Db 113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy 64 FDKTPEP 70
Db 173 FTKPPPP 179

RESULT 10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 58.6%; Score 212; DB 1; Length 355;
Best Local Similarity 55.7%; Pred. No. 1.1e-15;
Matches 39; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy 61 DRIFDKTPEP 70
Db 170 ERIFAKAPEP 179

RESULT 11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 57.7%; Score 209; DB 1; Length 358;
Best Local Similarity 52.9%; Pred. No. 2.3e-15;
Matches 37; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|| :| || |::||: |:|||| | : || ::|||: |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176
```

RESULT 12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.7%;  Score 209;  DB 2;  Length 359;
  Best Local Similarity   53.6%;  Pred. No. 2.4e-15;
  Matches   37;  Conservative   13;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|| :| || ||:||| |:|:|| || ||| :|:: |  |: |||::|||| 
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179
```

RESULT 13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.7%;  Score 209;  DB 2;  Length 385;
  Best Local Similarity   52.9%;  Pred. No. 2.5e-15;
  Matches   37;  Conservative   17;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|| :| || ||:||: |:|||||:| ::|||||::|: |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
           :|| : |:|
Db    197 QKIFQRPPDP 206
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             55.8%;  Score 202;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 1.4e-14;
  Matches   38;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| ||:|| :| || |::||| |||:|: ||  || |:||| |: :: ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             55.2%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 2.3e-14;
  Matches   38;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |:|| :| || |::||| |||||: || ||| |::|| |: |: ||||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171
```

Search completed: February 3, 2001, 02:17:37
Job time: 2209 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:38 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	215	59.4	359	3	US-08-809-103B-2	Sequence 2, Appli
2	215	59.4	359	3	US-08-809-103B-4	Sequence 4, Appli
3	215	59.4	359	3	US-08-809-103B-6	Sequence 6, Appli
4	215	59.4	359	3	US-08-809-103B-8	Sequence 8, Appli
5	68.5	18.9	1713	3	US-08-600-982-24	Sequence 24, Appl
6	68.5	18.9	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
7	65.5	18.1	2161	1	US-07-745-206A-2	Sequence 2, Appli
8	65.5	18.1	2161	1	US-08-455-543A-49	Sequence 49, Appl
9	65.5	18.1	2161	1	US-08-455-543A-51	Sequence 51, Appl
10	65.5	18.1	2161	2	US-08-223-305C-49	Sequence 49, Appl
11	65.5	18.1	2161	2	US-08-223-305C-51	Sequence 51, Appl
12	65.5	18.1	2161	2	US-08-311-363-2	Sequence 2, Appli
13	63.5	17.5	334	5	5290690-11	Patent No. 5290690
14	60	16.6	446	2	US-08-672-814D-11	Sequence 11, Appl
15	59	16.3	454	3	US-08-446-100-31	Sequence 31, Appl
16	58.5	16.2	439	1	US-07-637-870-9	Sequence 9, Appli
17	58.5	16.2	439	1	US-07-637-399-6	Sequence 6, Appli
18	58.5	16.2	439	1	US-08-112-703-6	Sequence 6, Appli
19	58	16.0	844	3	US-09-029-267-20	Sequence 20, Appl
20	57	15.7	454	3	US-08-446-100-30	Sequence 30, Appl
21	56	15.5	454	3	US-08-446-100-26	Sequence 26, Appl
22	56	15.5	454	3	US-08-446-100-27	Sequence 27, Appl
23	56	15.5	454	3	US-08-446-100-28	Sequence 28, Appl
24	56	15.5	454	3	US-08-446-100-29	Sequence 29, Appl
25	56	15.5	548	2	US-08-871-266B-17	Sequence 17, Appl
26	56	15.5	548	2	US-09-018-864A-17	Sequence 17, Appl
27	56	15.5	548	3	US-08-871-267B-23	Sequence 23, Appl
28	55.5	15.3	111	1	US-08-185-414E-2	Sequence 2, Appli
29	55.5	15.3	770	1	US-08-445-135-2	Sequence 2, Appli
30	54	14.9	485	1	US-07-881-075-1	Sequence 1, Appli
31	54	14.9	485	1	US-08-120-827-1	Sequence 1, Appli
32	54	14.9	485	1	US-08-478-675-1	Sequence 1, Appli
33	54	14.9	589	3	US-08-942-008-2	Sequence 2, Appli
34	53.5	14.8	103	1	US-08-552-142A-13	Sequence 13, Appl
35	53.5	14.8	497	2	US-08-570-311-2	Sequence 2, Appli
36	53.5	14.8	497	2	US-08-353-485-2	Sequence 2, Appli
37	53.5	14.8	2628	2	US-08-570-311-14	Sequence 14, Appl
38	53	14.6	751	3	US-09-036-987A-24	Sequence 24, Appl
39	53	14.6	783	5	5231168-2	Patent No. 5231168
40	52.5	14.5	268	2	US-08-997-080-186	Sequence 186, App
41	52.5	14.5	268	2	US-08-997-362-186	Sequence 186, App
42	52.5	14.5	268	3	US-09-095-855-186	Sequence 186, App
43	52.5	14.5	340	3	US-09-095-855-205	Sequence 205, App
44	52.5	14.5	439	1	US-07-762-681A-1	Sequence 1, Appli
45	52.5	14.5	921	1	US-08-188-582-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match           59.4%;  Score 215;  DB 3;  Length 359;
  Best Local Similarity  55.1%;  Pred. No. 1.1e-19;
  Matches  38;  Conservative  12;  Mismatches  19;  Indels  0;  Gaps  0;
```

```
Qy      2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             59.4%;  Score 215;  DB 3;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.1e-19;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             59.4%;  Score 215;  DB 3;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.1e-19;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             59.4%;  Score 215;  DB 3;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.1e-19;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             18.9%;  Score 68.5;  DB 3;  Length 1713;
  Best Local Similarity   36.1%;  Pred. No. 4.6;
  Matches   22;  Conservative   11;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGAAAAGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
             || | |       :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   6
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24
```

```
  Query Match             18.9%;  Score 68.5;  DB 4;  Length 1713;
  Best Local Similarity   36.1%;  Pred. No. 4.6;
  Matches   22;  Conservative   11;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGAAAAGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           || | |      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   7
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             18.1%;  Score 65.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGAAAAGGCQTS----------------------------NDAAAEALNASSK 34
           |      ||  | ||  :|                            | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   8
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             18.1%;  Score 65.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;
```

```
Qy       4 WGEFQVDGAAAAGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  9
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match             18.1%;  Score 65.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGAAAAGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  10
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
```

```
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             18.1%;  Score 65.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGAAAAGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  11
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             18.1%;  Score 65.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGAAAAGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  12
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,363
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-51506
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (619)238-0999
;       TELEFAX:  (619)238-0062
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2161 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             18.1%;  Score 65.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGAAAAGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  13
5290690-11
;Patent No. 5290690
;     APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSSENS, PATRICK
;MATTHYSSENS, GASTON;WODAK, SHOSHANA;QUAX, WILHELMUS J.
;     TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
;STABILITY OF PROTEINS
;     NUMBER OF SEQUENCES: 22
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/398,706
;       FILING DATE: 25-AUG-1989
;SEQ ID NO:11:
;       LENGTH: 334
5290690-11


  Query Match             17.5%;  Score 63.5;  DB 5;  Length 334;
  Best Local Similarity   28.4%;  Pred. No. 2.4;
  Matches   19;  Conservative   12;  Mismatches   27;  Indels    9;  Gaps    1;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKE---------EALQIIREKIPEKYLFQ 52
           | |||  ||    : |  |  : ||: | | :|:         | : ::     :||  :
Db      82 LAWGEIGVDIVVESTGRFTKREDAAKHLEAGAKKVIISAPAKVENITVVMGVNQDKYDPK 141

Qy      53 FHNLNSN 59
             |:: ||
Db     142 AHHVISN 148


RESULT  14
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
;  GENERAL INFORMATION:
;     APPLICANT:  Leung, David W.
;     APPLICANT:  Weeks, Reitha
;     TITLE OF INVENTION:  MAMMALIAN
;     NUMBER OF SEQUENCES:  13
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cell Therapeutics, Inc.
;       STREET:  201 Elliott Avenue West
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  U.S.A.
;       ZIP:  98119
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
;       COMPUTER:  PC Clone (486 microprocessor)
;       OPERATING SYSTEM:  MS-DOS Version 6.1, Windows NT
;       SOFTWARE:  WORD 6.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/672,814D
;       FILING DATE:  28-Jun-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Faciszewski, Stephen
;       REGISTRATION NUMBER:  36,131
;       REFERENCE/DOCKET NUMBER:  1803
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206)282-7100
;       TELEFAX:  (206)284-6206
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  446
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  no
;     ANTI-SENSE:  no
;     FRAGMENT TYPE:
;     ORIGINAL SOURCE:
;       ORGANISM:  Drosophila
;       STRAIN:
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:
;       HAPLOTYPE:
;       TISSUE TYPE:
;       CELL TYPE:
;       CELL LINE:
;       ORGANELLE:
;     FEATURE:
;       NAME/KEY:  CDP-diacylglycerol synthase (CDS)
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-08-672-814D-11


  Query Match             16.6%;  Score 60;  DB 2;  Length 446;
  Best Local Similarity   34.0%;  Pred. No. 9.9;
  Matches   17;  Conservative    9;  Mismatches   22;  Indels    2;  Gaps    1;

Qy      20 TSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
             : :||| :  :|:   : :   |||||:       |  || : |||||
Db      20 SGSDAANKRNSAADSSDHVDSEEEKIPEEKFVD--ELAKNLPQGTDKTPE 67


RESULT  15
US-08-446-100-31
; Sequence 31, Application US/08446100
; Patent No. 6001625
;  GENERAL INFORMATION:
;     APPLICANT:  Broomfield, Clarence A
;     APPLICANT:  Millard, Charles B
;     APPLICANT:  Lockridge, Oksana
;     TITLE OF INVENTION:  Site-Directed Mutagenesis of Esterases
;     NUMBER OF SEQUENCES:  31
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Hendricks and Assoc.
;       STREET:  9669 A Main Street, P.O. Box 2509
;       CITY:  Fairfax
;       STATE:  VA
;       COUNTRY:  US
;       ZIP:  22031
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/446,100
;       FILING DATE:  19-MAY-1995
```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenna
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: broomfield
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 425-4250
; TELEFAX: (703) 425-2767
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 454 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: human esterases
US-08-446-100-31

Query Match 16.3%; Score 59; DB 3; Length 454;
Best Local Similarity 31.8%; Pred. No. 14;
Matches 21; Conservative 10; Mismatches 25; Indels 10; Gaps 4;

```
Qy      13 AAAGGCQTSNDAA-AEALNASSKEEALQIIREKIPEKYLFQF-------HNLNSNLD-RI 63
           |   ||:|:  ||    |   ::|| |:    ||   ||: :       | | : :|  :
Db     156 AITAGCKTTTSAAMVHCLRQKTEEELLETTL-KIGNSYLWTYRETQRESHLLGTVIDGML 214

Qy      64 FDKTPE 69
             ||||
Db     215 LLKTPE 220
```

Search completed: February 3, 2001, 02:19:40
Job time: 1702 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:25 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-2
Perfect score: 362
Sequence: 1 TLVWGEFQVDGAAAAGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	215	59.4	353	18	W34338	Bean golden mosaic
2	215	59.4	353	18	W34332	Bean golden mosaic
3	215	59.4	353	18	W34333	Bean golden mosaic
4	215	59.4	353	18	W34334	Bean golden mosaic
5	215	59.4	353	18	W34335	Bean golden mosaic
6	215	59.4	359	17	R88870	Sardinian tomato y
7	215	59.4	359	17	R88871	Sardinian tomato y
8	215	59.4	359	17	R88872	Sardinian tomato y
9	213	58.8	353	8	P70407	ORF 4 gene product
10	213	58.8	361	18	W34336	Tomato mottle viru
11	213	58.8	361	18	W34324	Tomato mottle viru
12	213	58.8	361	18	W34325	Tomato mottle viru
13	213	58.8	361	18	W34326	Tomato mottle viru
14	206.5	57.0	361	8	P70562	Product of ORF 4 f
15	201	55.5	362	19	W56495	Tobacco leaf curli
16	200	55.2	357	18	W34329	Tomato yellow leaf
17	200	55.2	357	18	W34330	Tomato yellow leaf
18	200	55.2	357	18	W34331	Tomato yellow leaf
19	192	53.0	357	18	W34337	Tomato yellow leaf
20	69.5	19.2	512	19	W68473	HIV-1 strain YBF30
21	68.5	18.9	1713	16	R70148	Deduced sequence o
22	65.5	18.1	619	13	R27651	Human calcium chan
23	65.5	18.1	2161	14	R33545	Sequence of the al
24	65.5	18.1	2161	16	R71001	Human neuronal cal
25	65.5	18.1	2161	16	R71002	Human neuronal cal
26	65.5	18.1	2161	19	W63137	Human calcium chan
27	65.5	18.1	2161	19	W63149	Human calcium chan
28	65	18.0	292	21	Y92317	Geminivirus Rep C1
29	65	18.0	335	21	Y92318	Mastrevirus Rep in
30	64.5	17.8	299	18	W55450	H. pylori ORF 02ae
31	63.5	17.5	334	11	R03211	Amino acid sequenc
32	60.5	16.7	111	20	Y37827	Amino acid sequenc
33	60.5	16.7	204	20	W89814	Protein encoded by
34	60.5	16.7	299	20	W89829	Protein encoded by
35	60.5	16.7	299	20	W89849	Antigen from clust
36	59.5	16.4	673	21	Y58996	Rice sulfite reduc
37	59	16.3	454	21	Y49470	Human carboxyleste
38	58.5	16.2	299	19	W98322	H. pylori GHPO 136
39	58.5	16.2	299	19	W73034	Helicobacter pylor
40	58	16.0	844	18	W09773	Histidinol dehydro
41	57.5	15.9	266	19	W72937	Mycobacterium tube
42	57.5	15.9	266	20	Y21957	Amino acid sequenc
43	57	15.7	454	21	Y49500	Human carboxyleste
44	56.5	15.6	249	20	W89822	Protein encoded by
45	56.5	15.6	269	20	W89860	Antigen 1 from clu

ALIGNMENTS

RESULT 1
W34338
ID W34338 standard; Protein; 353 AA.
XX
AC W34338;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 protein.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93314.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

XX
PS Example 5; Page 100-102; 132pp; English.
XX
CC This sequence comprises the bean golden mosaic virus (BGMV) C1
CC protein that is required for replication. The invention involves
CC production of transgenic plants containing DNA comprising C1 or AC1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from BGMV, tomato mottle virus or tomato yellow leaf
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC that have mutations in the highly conserved DNA-nicking domain
CC and/or the NTP-binding domains.
XX
SQ Sequence 353 AA;

Query Match 59.4%; Score 215; DB 18; Length 353;
Best Local Similarity 57.1%; Pred. No. 1.4e-18;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
|: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db 110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy 61 DRIFDKTPEP 70
:||| | |||
Db 170 erifvkvpep 179

RESULT 2
W34332
ID W34332 standard; Protein; 353 AA.
XX
AC W34332;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93290.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 103-105; 132pp; English.
XX
CC This protein comprises a control mutant of the bean golden mosaic
CC virus (BGMV) C1 protein (see W34338) that is required for
CC replication. It is encoded by mutated C1 open reading frame
CC BGAC190 (see T93290). The invention involves production of
CC transgenic plants containing DNA comprising geminivirus C1 or AC1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from BGMV, tomato mottle virus or tomato yellow leaf
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC that have mutations in the highly conserved DNA-nicking and/or the
CC NTP-binding domains.
XX
SQ Sequence 353 AA;

Query Match 59.4%; Score 215; DB 18; Length 353;
Best Local Similarity 57.1%; Pred. No. 1.4e-18;
Matches 40; Conservative 14; Mismatches 16; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
|: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db 110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy 61 DRIFDKTPEP 70
:||| | |||
Db 170 erifvkvpep 179

RESULT 3
W34333
ID W34333 standard; Protein; 353 AA.
XX
AC W34333;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93291.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 107-109; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminiviral
CC replication during infection. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGMV,
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the

```
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            59.4%;  Score 215;  DB 18;  Length 353;
  Best Local Similarity  57.1%;  Pred. No. 1.4e-18;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179


RESULT   4
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            59.4%;  Score 215;  DB 18;  Length 353;
  Best Local Similarity  57.1%;  Pred. No. 1.4e-18;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179


RESULT   5
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            59.4%;  Score 215;  DB 18;  Length 353;
  Best Local Similarity  57.1%;  Pred. No. 1.4e-18;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169
```

```
Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             59.4%;  Score 215;  DB 17;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.4e-18;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   7
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             59.4%;  Score 215;  DB 17;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.4e-18;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   8
```

```
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             59.4%;  Score 215;  DB 17;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.4e-18;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|| :| || ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   9
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             58.8%;  Score 213;  DB 8;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 2.4e-18;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||| :| || |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93294.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 3.3; Page 57-58; 132pp; English.
XX
CC This protein comprises the wild-type AC1 protein of tomato mottle
CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1
CC gene (see T93294) must be expressed for efficient replication of
CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a
CC DNA binding site specific to the DNA-A common region, a DNA nicking
CC activity, and an NTP binding activity. The invention involves
CC production of transgenic plants containing DNA comprising AC1 or C1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC W34324-35) that have mutations in the highly conserved DNA-nicking
CC domain and/or the NTP-binding domain.
XX
SQ Sequence 361 AA;

Query Match 58.8%; Score 213; DB 18; Length 361;
Best Local Similarity 54.3%; Pred. No. 2.5e-18;
Matches 38; Conservative 17; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy 61 DRIFDKTPEP 70
Db 170 erifakapep 179

RESULT 11
W34324
ID W34324 standard; Protein; 361 AA.
XX
AC W34324;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93282.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 3.4; Page 60-62; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant, designated
CC ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC W34336). It is encoded by a mutant AC1 gene (see T93282) of
CC tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC NTP-binding domains. The AC1 gene (see also T93294) must be
CC expressed for efficient replication of the two genomic components,
CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC involves production of transgenic plants containing DNA comprising
CC geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC interfere in trans with geminiviral replication during infection.
CC Such transgenic plants are resistant to viral infection. The
CC AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC polypeptides (see W34324-35) that have mutations in the highly
CC conserved DNA-nicking and/or NTP-binding domains.
XX
SQ Sequence 361 AA;

Query Match 58.8%; Score 213; DB 18; Length 361;
Best Local Similarity 54.3%; Pred. No. 2.5e-18;
Matches 38; Conservative 17; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
Db 110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy 61 DRIFDKTPEP 70
Db 170 erifakapep 179

RESULT 12
W34325
ID W34325 standard; Protein; 361 AA.
XX
AC W34325;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93283.

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336). It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain. The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection. The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             58.8%;  Score 213;  DB 18;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 2.5e-18;
  Matches   38;  Conservative   17;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|| :| || |::||: |:|||||| : || ::||: |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336). It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain. The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection. The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             58.8%;  Score 213;  DB 18;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 2.5e-18;
  Matches   38;  Conservative   17;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|| :| || |::||: |:|||||| : || ::||: |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             57.0%;  Score 206.5;  DB 8;  Length 361;
  Best Local Similarity   63.1%;  Pred. No. 1.6e-17;
  Matches   41;  Conservative   10;  Mismatches   13;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           | ||||| :| || |::||| ||||||: || :|| |:||| |::|: |||||| || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             55.5%;  Score 201;  DB 19;  Length 362;
  Best Local Similarity   50.0%;  Pred. No. 7.4e-17;
  Matches   42;  Conservative   11;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVDGAAAAGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | :| ||||| :|   ||::||| |||:|: ||  || |:||| |: :: ||||||||||
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 RI---------------FDKTPE 69
           ||               ||: ||
Db     171 rifapplevfvcpfssssfdqvpe 194



Search completed: February  3, 2001, 02:15:26
Job time: 3433 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:50 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	276	77.1	226	12	O09727	O09727 leonurus mo
2	271	75.7	364	12	Q9QS55	Q9qs55 sweet potat
3	264	73.7	226	12	Q9WHF6	Q9whf6 tomato mild
4	262	73.2	361	12	Q67574	Q67574 bean golden
5	260	72.6	149	12	P88975	P88975 macroptiliu
6	260	72.6	225	12	Q9QDB1	Q9qdb1 cowpea gold
7	260	72.6	233	12	Q9YLA4	Q9yla4 macroptiliu
8	259	72.3	185	12	Q98693	Q98693 sida golden
9	251	70.1	190	12	Q9Z089	Q9z089 tobacco lea
10	251	70.1	190	12	Q9Z084	Q9z084 tobacco lea
11	250	69.8	190	12	Q9W827	Q9w827 tobacco lea
12	250	69.8	208	12	Q9Z0C4	Q9z0c4 tobacco lea
13	250	69.8	208	12	Q9Z0C0	Q9z0c0 tobacco lea
14	250	69.8	208	12	Q9Z0B8	Q9z0b8 tobacco lea
15	249	69.6	208	12	Q9Z0A0	Q9z0a0 tobacco lea
16	248	69.3	208	12	Q9Z0C6	Q9z0c6 tobacco lea
17	247	69.0	362	12	O56816	O56816 chayote mos
18	246	68.7	208	12	Q9Z0B6	Q9z0b6 tobacco lea
19	245	68.4	203	12	Q9Z083	Q9z083 tobacco lea
20	245	68.4	363	12	O73577	O73577 cotton leaf
21	244	68.2	362	12	Q9IN48	Q9in48 okra enatio
22	243	67.9	234	12	O39180	O39180 potato yell
23	242	67.6	231	12	Q96620	Q96620 african tom
24	242	67.6	353	12	O72692	O72692 beet curly
25	242	67.6	354	12	Q65438	Q65438 beet curly
26	242	67.6	359	12	Q9YZV4	Q9yzv4 tomato yell
27	242	67.6	359	12	Q9YZV2	Q9yzv2 tomato yell
28	242	67.6	359	12	Q9YUX7	Q9yux7 tomato yell
29	242	67.6	359	12	Q9YL27	Q9yl27 tomato yell
30	241	67.3	358	12	Q65418	Q65418 beet curly
31	241	67.3	359	12	Q88942	Q88942 tomato yell
32	239	66.8	307	12	Q9IET7	Q9iet7 cotton leaf
33	239	66.8	361	12	O72723	O72723 cotton leaf
34	238	66.5	190	12	Q9Z0A7	Q9z0a7 tobacco lea
35	237	66.2	190	12	Q9Z086	Q9z086 tobacco lea
36	234	65.4	363	12	O72705	O72705 cotton leaf
37	234	65.4	363	12	O72719	O72719 cotton leaf
38	233	65.1	362	12	Q9YYT7	Q9yyt7 althea rose
39	233	65.1	363	12	O73494	O73494 okra yellow
40	231	64.5	306	12	Q9IET1	Q9iet1 cotton leaf
41	231	64.5	354	12	Q9IN42	Q9in42 south afric
42	230	64.2	351	12	Q9Q9R3	Q9q9r3 dicliptera
43	230	64.2	358	12	Q9JEA2	Q9jea2 cassava gem
44	230	64.2	360	12	O72707	O72707 cotton leaf
45	230	64.2	360	12	O72710	O72710 cotton leaf

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.1%; Score 276; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 3.1e-24;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||||||||:||| || ||||| || |     |||||:||:||||||||||||||
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180
```

RESULT 2
Q9QS55

```
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.7%;  Score 271;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 2e-23;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||||||| ||:||||| || | |  |||||||||:|||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   3
Q9WHF6
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             73.7%;  Score 264;  DB 12;  Length 226;
  Best Local Similarity   70.0%;  Pred. No. 7.5e-23;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||||||||||| ||:||||| || | |  |||||||||:|||||||||| |||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   4
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             73.2%;  Score 262;  DB 12;  Length 361;
  Best Local Similarity   73.1%;  Pred. No. 2.1e-22;
  Matches   49;  Conservative    9;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || |||||||||| ||:||||: || |||   |:|||:||:|||:||:|:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   5
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
```

```
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             72.6%;  Score 260;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 1.4e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||||| || : :  ||::|::||:|||:|||:|||:|||
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    112 DRIFMKDPEP 121


RESULT   6
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             72.6%;  Score 260;  DB 12;  Length 225;
  Best Local Similarity   73.1%;  Pred. No. 2.2e-22;
  Matches   49;  Conservative    8;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
          || ||:|||||||| || |||| || |||   |:|||:||:|||:|||:|||:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT   7
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             72.6%;  Score 260;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 2.2e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||||| || : :  ||::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   8
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             72.3%;  Score 259;  DB 12;  Length 185;
  Best Local Similarity   68.6%;  Pred. No. 2.3e-22;
```

```
  Matches   48;  Conservative   11;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||:||||| || : :   ||:|||||:||:||||:|||:||:
Db     89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    149 DRIFSKPPEP 158


RESULT   9
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             70.1%;  Score 251;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:||||||||| :||| | || ||| | || |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  10
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             70.1%;  Score 251;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:||||||||| :||| | || ||| | || |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             69.8%;  Score 250;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.6e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          || || |||||||||||||| :||| | || |:| | || |||||:|: ::||:|||||||
Db     85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  12
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
```

```
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             69.8%;  Score 250;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 2.8e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| | || |:| | || ||||||:|: ::||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  13
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             69.8%;  Score 250;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 2.8e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| | || ||| | || |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  14
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;


  Query Match             69.8%;  Score 250;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 2.8e-21;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| | || ||| | || |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0A0
ID   Q9Z0A0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0A0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE MINO2-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MINO2;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001308; BAA34020.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
```

```
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23423 MW;  D56D242F1BBEC55D CRC64;


  Query Match             69.6%;  Score 249;  DB 12;  Length 208;
  Best Local Similarity   57.6%;  Pred. No. 3.7e-21;
  Matches   49;  Conservative    8;  Mismatches   12;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| | || | | |||| |||||:|: ::| :|||||||
Db     102 TLEWGSFQIDGRSARGGCQNANDACAEALNAGSKAAALSIIREKLPKDFIFNYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


Search completed: February  3, 2001, 02:22:50
Job time: 1887 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:24 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	332	92.7	352	1	VAL1_TGMV	P03567 tomato gold
2	243	67.9	361	1	VAL1_PYMVV	P27258 potato yell
3	230	64.2	358	1	VAL1_CLVK	P14982 cassava lat
4	230	64.2	358	1	VAL1_CLVN	P14972 cassava lat
5	222	62.0	362	1	VAL1_TYLCA	P36279 tomato yell
6	217	60.6	355	1	VAL1_ABMVW	P21947 abutilon mo
7	216	60.3	349	1	VAL1_PHUV	Q06923 pepper huas
8	216	60.3	359	1	VAL1_TYLCU	P38609 tomato yell
9	215	60.1	353	1	VAL1_BGMV	P05175 bean golden
10	214	59.8	361	1	VAL1_TMOV	Q06657 tomato mott
11	212	59.2	359	1	VAL1_TYLCM	P27260 tomato yell
12	208	58.1	358	1	VAL1_BCTV	P14991 beet curly
13	190	53.1	357	1	VAL1_TYLCV	P27259 tomato yell
14	115	32.1	347	1	VAL1_SLCV	P29048 squash leaf
15	69	19.3	577	1	T2FA_DROME	Q05913 drosophila
16	66	18.4	207	1	IL6_MARMO	O35736 marmota mon
17	66	18.4	508	1	CROC_DROME	P32027 drosophila
18	65	18.2	671	1	HMOC_DROME	P22810 drosophila
19	62.5	17.5	1077	1	HLES_DROME	Q02308 drosophila
20	60.5	16.9	829	1	E74A_DROME	P20105 drosophila
21	60.5	16.9	883	1	E74B_DROME	P11536 drosophila
22	60	16.8	314	1	PMXB_HUMAN	Q99453 homo sapien
23	60	16.8	314	1	PMXB_MOUSE	O35690 mus musculu
24	59.5	16.6	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
25	59	16.5	379	1	CYB_ASTPE	Q33818 asterina pe
26	59	16.5	498	1	CALR_PIG	P25117 sus scrofa
27	59	16.5	513	1	HEM0_CHICK	P18080 gallus gall
28	59	16.5	3591	1	FHAB_BORPE	P12255 bordetella
29	58.5	16.3	1264	1	MOLR_ECOLI	P33345 escherichia
30	58.5	16.3	1713	1	LMA3_HUMAN	Q16787 homo sapien
31	58	16.2	732	1	CYG4_HUMAN	P33402 homo sapien
32	58	16.2	1081	1	UL52_HSVEB	P28962 equine herp
33	58	16.2	1992	1	TR12_HUMAN	Q14669 homo sapien
34	57.5	16.1	303	1	ALB3_MAIZE	P10593 zea mays (m
35	57.5	16.1	557	1	SASB_ANAPL	Q04791 anas platyr
36	57	15.9	387	1	VE2_HPV41	P27552 human papil
37	57	15.9	401	1	HME1_MOUSE	P09065 mus musculu
38	57	15.9	3080	1	POLG_ZYMVC	P18479 z genome po
39	56.5	15.8	664	1	PLB1_YEAST	P39105 saccharomyc
40	56.5	15.8	1205	1	NKC1_MOUSE	P55012 mus musculu
41	56	15.6	105	1	RLA2_LEIBR	O44010 leishmania
42	56	15.6	398	1	ODO2_AZOVI	P20708 azotobacter
43	56	15.6	447	1	CDSA_DROME	P56079 d phosphati
44	56	15.6	582	1	HEM0_OPSTA	P43090 opsanus tau
45	56	15.6	1114	1	DEFI_CHICK	P36197 gallus gall

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 92.7%; Score 332; DB 1; Length 352;
Best Local Similarity 92.9%; Pred. No. 6.5e-32;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||| || |||  ||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             67.9%;  Score 243;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 2.2e-21;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || ||||| || : :  ||::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             64.2%;  Score 230;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 7.3e-20;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| | || : | : || :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             64.2%;  Score 230;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 7.3e-20;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| | || : | : || :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             62.0%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 6.5e-19;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||| | ||   | : || ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             60.6%;  Score 217;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 2.5e-18;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| |+:+|: | || |    :|| |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
```

```
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.3%;  Score 216;  DB 1;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 3.2e-18;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::|| | || ::|| ||||||:|: |: :  |||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             60.3%;  Score 216;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 3.3e-18;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.1%;  Score 215;  DB 1;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 4.3e-18;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:|||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             59.8%;  Score 214;  DB 1;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.7e-18;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: | || | |  :|| ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             59.2%;  Score 212;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 9.8e-18;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  12
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             58.1%;  Score 208;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 2.9e-17;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: | || |:|   ||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             53.1%;  Score 190;  DB 1;  Length 357;
  Best Local Similarity   60.7%;  Pred. No. 3.8e-15;
  Matches   37;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |:||||||||| |::||| | || : | : || |::|| |: |: ||||||||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             32.1%;  Score 115;  DB 1;  Length 347;
  Best Local Similarity   36.4%;  Pred. No. 2.7e-06;
  Matches   24;  Conservative   14;  Mismatches   24;  Indels    4;  Gaps    1;

Qy      5 GEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRIF 64
          |:::| |    |     :|    |: | ||  || ||:   |: :: ':||| :|::|:|
Db    116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy     65 DKTPEP 70
          | |||
Db    172 QKPPEP 177


RESULT  15
T2FA_DROME
ID   T2FA_DROME     STANDARD;      PRT;   577 AA.
AC   Q05913;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   TRANSCRIPTION INITIATION FACTOR IIF, ALPHA SUBUNIT (TFIIF-ALPHA)
DE   (TRANSCRIPTION FACTOR 5, LARGE CHAIN) (TF5A).
GN   TFIIF-ALPHA.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93219133; PubMed=8464745;
RA   Gong D.-W., Horikoshi M., Nakatani Y.;
RT   "Analysis of cDNA encoding Drosophila transcription initiation factor
RT   TFIIF alpha (RAP74).";
RL   Nucleic Acids Res. 21:1492-1492(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93219101; PubMed=8464716;
RA   Kephart D., Price M.P., Burton Z.F., Finkelstein A.,
RA   Greenblalt J., Price D.H.;
RT   "Cloning of a Drosophila cDNA with sequence similarity to human
RT   transcription factor RAP74.";
RL   Nucleic Acids Res. 21:1319-1319(1993).
RN   [3]
RP   FUNCTION.
RX   MEDLINE=94230464; PubMed=8175788;
RA   Kephart D., Wang B.Q., Burton Z.F., Price D.H.;
RT   "Functional analysis of Drosophila factor 5 (TFIIF), a general
RT   transcription factor.";
RL   J. Biol. Chem. 269:13536-13543(1994).
CC   -!- FUNCTION: TFIIF IS A GENERAL TRANSCRIPTION INITIATION FACTOR THAT
CC       BINDS TO RNA POLYMERASE II AND HELPS TO RECRUIT IT TO THE
CC       INITIATION COMPLEX IN COLLABORATION WITH TFIIB. IT PROMOTES
CC       TRANSCRIPTION ELONGATION.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- SIMILARITY: TO OTHER TRANSCRIPTION FACTOR IIF, ALPHA SUBUNIT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L10331; -; NOT_ANNOTATED_CDS.
DR   EMBL; X66982; CAA47391.1; -.
DR   PIR; S30237; S30237.
DR   TRANSFAC; T02173; -.
DR   FLYBASE; FBgn0010282; TfIIF-alpha.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   CONFLICT     20     20       R -> A (IN REF. 2).
FT   CONFLICT     76     76       G -> E (IN REF. 2).
FT   CONFLICT    528    528       A -> P (IN REF. 2).
SQ   SEQUENCE   577 AA;  64453 MW;  86EE867397F505B3 CRC64;


  Query Match             19.3%;  Score 69;  DB 1;  Length 577;
  Best Local Similarity   36.4%;  Pred. No. 1.3;
  Matches   20;  Conservative    9;  Mismatches   20;  Indels    6;  Gaps    1;

Qy     13 SARGGCQTSNDAAAAALAASSAAAALQIIRE------KIPEKYLFQFHNLNSNLD 61
          ||  |  ||  |||||  || :|::   ::|      |:|:|:     |   |:|
Db     10 SAASGSSTSARAAAAASVASGSASSSANVQEFKIRVPKMPKKHHVMRFNATLNVD 64


Search completed: February  3, 2001, 02:24:25
Job time: 637 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:37 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	332	92.7	352	1	QQCVL1	AL1 protein - toma
2	243	67.9	361	1	QQCVPT	AL1 protein - pota
3	230	64.2	358	2	S07594	hypothetical prote
4	222	62.0	362	1	JQ1887	AL1 protein - toma
5	217	60.6	355	1	QQCVW1	AV1 protein - abut
6	216	60.3	349	2	JQ2300	replicase - pepper
7	216	60.3	349	2	S31875	AL1 protein - pepp
8	216	60.3	359	2	S39211	gene C1 protein -
9	214	59.8	358	1	JQ1870	AL1 protein - toma
10	213	59.5	351	2	JQ2327	AL1 protein - Indi
11	212	59.2	359	2	S22593	hypothetical prote
12	208	58.1	385	2	S28360	AL1 protein - beet
13	207	57.8	360	2	S59885	replication-associ
14	203	56.7	359	2	S39235	gene C1 protein -
15	190	53.1	357	1	QQCVC1	AL1 protein - toma
16	120	33.5	131	2	S45059	AC1 protein (clone
17	115	32.1	347	1	QQCVS1	AL1 protein - squa
18	69	19.3	577	2	S30237	transcription init
19	66	18.4	508	2	S59870	fork head domain p
20	66	18.4	840	2	T36175	probable large ATP
21	65	18.2	671	2	A35912	homeotic protein o
22	63	17.6	392	2	T44362	acetyl-CoA C-acety
23	62.5	17.5	700	2	T05841	spliceosome-associ
24	62.5	17.5	1077	2	A44067	serine-rich protei
25	62	17.3	409	2	T16781	hypothetical prote
26	61.5	17.2	327	2	F83408	hypothetical prote
27	61.5	17.2	705	2	T31157	hypothetical prote
28	61.5	17.2	1541	2	T02831	AAA protein L4171.
29	61	17.0	339	2	T26328	hypothetical prote
30	61	17.0	371	2	B69502	alcohol dehydrogen
31	61	17.0	1265	2	T47626	structural mainten
32	61	17.0	1510	2	T33100	hypothetical prote
33	60.5	16.9	509	2	T02942	O-succinylhomoseri
34	60.5	16.9	829	2	A34692	ecdysone-induced p
35	60.5	16.9	883	2	S04722	puff 74E protein -
36	60	16.8	314	2	JC5273	paired type homeob
37	59.5	16.6	201	2	A81380	probable flagellar
38	59.5	16.6	470	2	T46814	gamma-aminobutyrat
39	59.5	16.6	481	2	A70091	probable phosphoes
40	59.5	16.6	509	2	T02940	O-succinylhomoseri
41	59	16.5	380	2	S70594	ubiquinol--cytochr
42	59	16.5	482	2	A39285	calcitonin recepto
43	59	16.5	498	2	I47130	calcitonin recepto
44	59	16.5	513	1	SYCHLE	5-aminolevulinate
45	59	16.5	3591	1	S21010	filamentous hemagg

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 92.7%; Score 332; DB 1; Length 352;
Best Local Similarity 92.9%; Pred. No. 1.1e-30;
Matches 65; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||||| || |||   ||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 67.9%; Score 243; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 1.9e-20;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:||||||| || ||||| || : :  ||::||:||:|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178
```

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 64.2%; Score 230; DB 2; Length 358;
Best Local Similarity 60.0%; Pred. No. 6.1e-19;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| | || : | : || :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178
```

RESULT 4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.0%; Score 222; DB 1; Length 362;
Best Local Similarity 52.9%; Pred. No. 5.1e-18;
Matches 45; Conservative 9; Mismatches 15; Indels 16; Gaps 1;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          || ||||:|||||||| |::||| | ||   | : || ::||  |: |: ||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194
```

RESULT 5
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.6%; Score 217; DB 1; Length 355;
Best Local Similarity 58.6%; Pred. No. 1.9e-17;
Matches 41; Conservative 11; Mismatches 18; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |  |||||:|||||||| ||:||: | || |    :|| |::|: |: |: | ||: |||
Db    110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179
```

RESULT 6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.3%; Score 216; DB 2; Length 349;
Best Local Similarity 58.6%; Pred. No. 2.4e-17;
Matches 41; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||||||:|||||||| |::||  | || ::||  |||||:|: |: :  |||| ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179
```

RESULT 7
S31875
AL1 protein - pepper rizado amarillo virus

```
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.3%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 2.4e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| |::||  | || ::||  |||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from the
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.3%;  Score 216;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.5e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   9
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 214;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 4.2e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: | || | |  :|| ::||: |: :: | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176


RESULT  10
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.5%;  Score 213;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 5.4e-17;
  Matches   41;  Conservative   10;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| |||| : | : ||:|:||  |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  11
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn,
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopa
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.2%;  Score 212;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 7.2e-17;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  12
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.1%;  Score 208;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 2.3e-16;
  Matches   39;  Conservative   13;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| ||:||: | || |:|   ||||::|: |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
           :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  13
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different countri
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             57.8%;  Score 207;  DB 2;  Length 360;
  Best Local Similarity   60.6%;  Pred. No. 2.7e-16;
  Matches   40;  Conservative   12;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| ||:|||||||| |::||| | |: : | |||| |:||| |: :: ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178


RESULT  14
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.7%;  Score 203;  DB 2;  Length 359;
  Best Local Similarity   53.6%;  Pred. No. 7.9e-16;
  Matches   37;  Conservative   12;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| | |: | | : || :|::  |  |:  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             53.1%;  Score 190;  DB 1;  Length 357;
  Best Local Similarity   60.7%;  Pred. No. 2.5e-14;
  Matches   37;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |:|||||||| |::||| | || : | : || |::|| |: |: |||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


Search completed: February  3, 2001, 02:17:38
Job time: 2210 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:40 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	212	59.2	359	3	US-08-809-103B-2	Sequence 2, Appli
2	212	59.2	359	3	US-08-809-103B-4	Sequence 4, Appli
3	212	59.2	359	3	US-08-809-103B-6	Sequence 6, Appli
4	212	59.2	359	3	US-08-809-103B-8	Sequence 8, Appli
5	63	17.6	103	1	US-08-552-142A-13	Sequence 13, Appl
6	60.5	16.9	509	2	US-08-737-524B-27	Sequence 27, Appl
7	59	16.5	482	1	US-07-792-885A-1	Sequence 1, Appli
8	59	16.5	482	1	US-08-142-439A-7	Sequence 7, Appli
9	59	16.5	482	2	US-08-869-477-7	Sequence 7, Appli
10	58.5	16.3	1713	3	US-08-600-982-24	Sequence 24, Appl
11	58.5	16.3	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
12	56.5	15.8	664	3	US-09-295-186-17	Sequence 17, Appl
13	56	15.6	446	2	US-08-672-814D-11	Sequence 11, Appl
14	54.5	15.2	485	1	US-07-881-075-1	Sequence 1, Appli
15	54.5	15.2	485	1	US-08-120-827-1	Sequence 1, Appli
16	54.5	15.2	485	1	US-08-478-675-1	Sequence 1, Appli
17	54.5	15.2	738	3	US-08-864-038A-3	Sequence 3, Appli
18	54	15.1	171	2	US-08-609-049A-22	Sequence 22, Appl
19	54	15.1	587	2	US-08-871-266B-18	Sequence 18, Appl
20	54	15.1	587	2	US-09-018-864A-18	Sequence 18, Appl
21	54	15.1	587	3	US-08-871-267B-24	Sequence 24, Appl
22	54	15.1	1345	2	US-08-977-767-3	Sequence 3, Appli
23	53.5	14.9	2100	2	US-08-808-793-23	Sequence 23, Appl
24	53.5	14.9	2100	3	US-08-772-512A-19	Sequence 19, Appl
25	52.5	14.7	440	3	US-09-100-664A-2	Sequence 2, Appli
26	52.5	14.7	440	3	US-09-100-664A-3	Sequence 3, Appli
27	52.5	14.7	440	3	US-09-100-664A-4	Sequence 4, Appli
28	52.5	14.7	513	3	US-09-100-193-3	Sequence 3, Appli
29	52.5	14.7	1184	2	US-08-918-914-1	Sequence 1, Appli
30	52.5	14.7	1184	3	US-08-996-083-3	Sequence 3, Appli
31	52	14.5	391	5	5229279-2	Patent No. 5229279
32	52	14.5	391	5	5512669-2	Patent No. 5512669
33	52	14.5	548	2	US-08-871-266B-17	Sequence 17, Appl
34	52	14.5	548	2	US-09-018-864A-17	Sequence 17, Appl
35	52	14.5	548	3	US-08-871-267B-23	Sequence 23, Appl
36	52	14.5	1958	1	US-07-945-283-2	Sequence 2, Appli
37	51.5	14.4	37	2	US-08-117-952-783	Sequence 783, App
38	51.5	14.4	303	3	US-09-045-973-1	Sequence 1, Appli
39	51.5	14.4	770	1	US-08-445-135-2	Sequence 2, Appli
40	51	14.2	439	1	US-07-762-681A-1	Sequence 1, Appli
41	51	14.2	1317	3	US-09-083-521-7	Sequence 7, Appli
42	50	14.0	76	5	5496550-6	Patent No. 5496550
43	50	14.0	98	2	US-07-814-220-2	Sequence 2, Appli
44	50	14.0	98	2	US-07-812-421-2	Sequence 2, Appli
45	50	14.0	126	1	US-08-083-949-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match            59.2%;  Score 212;  DB 3;  Length 359;
  Best Local Similarity  56.5%;  Pred. No. 1.2e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             59.2%;  Score 212;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  3
US-08-809-103B-6
```

```
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             59.2%;  Score 212;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             59.2%;  Score 212;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-552-142A-13
; Sequence 13, Application US/08552142A
; Patent No. 5695995
;  GENERAL INFORMATION:
;    APPLICANT:  Weintraub, Harold M.
;    APPLICANT:  Lee, Jacqueline E.
;    APPLICANT:  Tapscott, Stephen J.
;    APPLICANT:  Hollenberg, Stanley M.
;    TITLE OF INVENTION:  Neurogenic Differentiation (NeuroD) Genes
;    TITLE OF INVENTION:  and Proteins
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen O'Connor Johnson KindnessPLLC
;      STREET:  1420 Fifth Avenue, Suite 2800
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-2347
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/552,142A
;      FILING DATE:  02-NOV-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/239,238
;      FILING DATE:  06-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US95/05741
;      FILING DATE:  08-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Broderick, Thomas F.
;      REGISTRATION NUMBER:  31,332
;      REFERENCE/DOCKET NUMBER:  FHCR-1-8933
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-682-8100
;      TELEFAX:  206-225-0709
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  103 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
US-08-552-142A-13


  Query Match             17.6%;  Score 63;  DB 1;  Length 103;
  Best Local Similarity   33.3%;  Pred. No. 0.28;
  Matches   19;  Conservative   11;  Mismatches   25;  Indels    2;  Gaps    1;

Qy      14 ARGGCQTSNDAAAAALAA--SSAAAALQIIREKIPEKYLFQFHNLNSNLDRIFDKTP 68
           :||  : : :||||   :   :|| | |  | |  ::   : ||||: || :   |
Db       6 SRGTGRRAGEAAAAGRXSLRGAAAXAAQERRVKANDRERNRMHNLNAALDALRSVLP 62


RESULT   6
US-08-737-524B-27
; Sequence 27, Application US/08737524B
; Patent No. 5912414
;  GENERAL INFORMATION:
;    APPLICANT:  CARL SAVERIO FALCO
;    APPLICANT:  DOMINICK ANTHONY GUIDA, JR.
;    APPLICANT:  MARY ELIZABETH HARNETT LOCKE
;    TITLE OF INVENTION:  NUCLEIC ACID FRAGMENTS, CHIMERIC
;    TITLE OF INVENTION:  GENES AND METHODS FOR INCREASING
;    TITLE OF INVENTION:  THE METHIONINE CONTENT OF THE SEEDS
;    TITLE OF INVENTION:  OF PLANTS
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  E. I. DU PONT DE NEMOURS AND COMPANY
;      STREET:  1007 MARKET STREET
;      CITY:  WILMINGTON
;      STATE:  DELAWARE
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  19898
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.50 INCH
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  MICROSOFT WINDOWS 95
;      SOFTWARE:  MICROSOFT WORD FOR WINDOWS 95 (7.0)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/737,524B
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LYNNE M. CHRISTENBURY
;      REGISTRATION NUMBER:  30,971
;      REFERENCE/DOCKET NUMBER:  BB-1059-A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  302-992-5481
;      TELEFAX:  302-773-0164
;      TELEX:  835420
;  INFORMATION FOR SEQ ID NO:  27:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  509 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-737-524B-27


  Query Match            16.9%;  Score 60.5;  DB 2;  Length 509;
  Best Local Similarity  54.8%;  Pred. No. 5;
  Matches   17;  Conservative    5;  Mismatches    8;  Indels    1;  Gaps    1;

Qy      11 GRSARGGCQTSNDAAAAALAASSAAAALQII 41
           || || |  :|: |||:| ||:|||| :  |
Db      71 GRRAR-GVASSHAAAASAAAAASAAAEVSAI 100


RESULT   7
US-07-792-885A-1
; Sequence 1, Application US/07792885A
; Patent No. 5516651
;  GENERAL INFORMATION:
;    APPLICANT:  Goldring, Steven R.
;    APPLICANT:  Gorn, Alan H.
;    APPLICANT:  Lin, Herb Y.
;    TITLE OF INVENTION:  MAMMALIAN CALCITONIN RECEPTOR
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/792,885A
;      FILING DATE:  19911115
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00786/074001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  482
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  N/A
;      TOPOLOGY:  N/A
US-07-792-885A-1


  Query Match            16.5%;  Score 59;  DB 1;  Length 482;
  Best Local Similarity  32.4%;  Pred. No. 7.4;
  Matches   12;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQI 40
           | ::|    : |   : :| ||| | ||:: |  ::|
Db     407 WNQYQAQRWAGRRSTRAANAAAATAAAAAALAETVEI 443


RESULT   8
US-08-142-439A-7
; Sequence 7, Application US/08142439A
; Patent No. 5670360
;  GENERAL INFORMATION:
;    APPLICANT:  Thorens, Bernard
;    TITLE OF INVENTION:  Receptor for the Glucagon-Like-Peptide-1
;    TITLE OF INVENTION:  (GLP-1)
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5670360o No. 5670360disk of No. 5670360th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10174-6201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/142,439A
;      FILING DATE:  24-NOV-93
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DK 398/92
;      FILING DATE:  25-MAR-92
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP93/00697
;      FILING DATE:  23-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Harrington, James J.
;      REGISTRATION NUMBER:  38,711
;      REFERENCE/DOCKET NUMBER:  3756.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 867 0123
;      TELEFAX:  212 867 0298
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  482 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Sus scrofa
US-08-142-439A-7


  Query Match            16.5%;  Score 59;  DB 1;  Length 482;
  Best Local Similarity  32.4%;  Pred. No. 7.4;
  Matches   12;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQI 40
           | ::|    : |   : :| ||| | ||:: |  ::|
Db     407 WNQYQAQRWAGRRSTRAANAAAATAAAAAALAETVEI 443


RESULT   9
US-08-869-477-7
; Sequence 7, Application US/08869477
; Patent No. 5846747
;  GENERAL INFORMATION:
;    APPLICANT:  Thorens, Bernard
;    TITLE OF INVENTION:  Receptor for the Glucagon-Like-Peptide-1
;    TITLE OF INVENTION:  (GLP-1)
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5846747o No. 5846747disk of No. 5846747th America, Inc.
;      STREET:  405 Lexington Avenue, Suite 6400
```

```
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  U.S.A.
;       ZIP:  10174-6201
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/869,477
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/142,439
;       FILING DATE:  24-NOV-93
;       APPLICATION NUMBER:  DK 398/92
;       FILING DATE:  25-MAR-92
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/EP93/00697
;       FILING DATE:  23-MAR-93
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Harrington, James J.
;       REGISTRATION NUMBER:  38,711
;       REFERENCE/DOCKET NUMBER:  3756.204-US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212 867 0123
;       TELEFAX:  212 867 0298
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  482 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Sus scrofa
US-08-869-477-7


  Query Match             16.5%;  Score 59;  DB 2;  Length 482;
  Best Local Similarity   32.4%;  Pred. No. 7.4;
  Matches   12;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQI 40
           | ::|    : |   : :| ||| | ||:: |  ::|
Db     407 WNQYQAQRWAGRRSTRAANAAAATAAAAAALAETVEI 443


RESULT  10
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;     APPLICANT:  Carter, William G.
;     APPLICANT:  Gil, Susanna A.
;     APPLICANT:  Ryan, Maureen C.
;     TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;     TITLE OF INVENTION:  Integrins
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;       STREET:  1420 Fifth Avenue
;       CITY:  Seattle
;       STATE:  WA
;       COUNTRY:  USA
;       ZIP:  98101-8100
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/600,982
;       FILING DATE:  02-SEP-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Shelton, Dennis K.
;       REGISTRATION NUMBER:  26,997
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 682-8100
;       TELEFAX:  (206) 224-0779
;   INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1713 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;       DESCRIPTION:  E170 protein as translated from sequence
;       DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;       DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             16.3%;  Score 58.5;  DB 3;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 46;
  Matches   21;  Conservative   13;  Mismatches   20;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAA--AAALAASSAAAALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  ||  ||  |||:: :||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  11
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;     APPLICANT:  Carter, William G.
;     APPLICANT:  Gil, Susanna A.
;     APPLICANT:  Ryan, Maureen C.
;     TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;     TITLE OF INVENTION:  Integrins
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;       STREET:  1420 Fifth Avenue
;       CITY:  Seattle
;       STATE:  WA
;       COUNTRY:  USA
;       ZIP:  98101-8100
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US94/10261A
;       FILING DATE:  02-SEP-1994
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Shelton, Dennis K.
;       REGISTRATION NUMBER:  26,997
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 682-8100
;       TELEFAX:  (206) 224-0779
;   INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1713 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24


  Query Match             16.3%;  Score 58.5;  DB 4;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 46;
  Matches   21;  Conservative   13;  Mismatches   20;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAA--AAALAASSAAAALQ-IIREKIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  ||  ||  |||:: :||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  12
US-09-295-186-17
; Sequence 17, Application US/09295186B
; Patent No. 6127137
; GENERAL INFORMATION:
;  APPLICANT:  Hasida, Miyoko
;  APPLICANT:  Tsutsumi, No. 6127137iko
;  APPLICANT:  Halkier, Torben
;  APPLICANT:  Stringer, Mary Ann
;  TITLE OF INVENTION:  An Acidic Phospholipase, Production, and
;  TITLE OF INVENTION:  Methods of Using Thereof (As Amended)
;  FILE REFERENCE:  4953.204-US
;  CURRENT APPLICATION NUMBER: US/09/295,186B
;  CURRENT FILING DATE:  1999-04-20
;  PRIOR APPLICATION NUMBER:  1215/96
;  PRIOR FILING DATE:  1996-10-31
;  PRIOR APPLICATION NUMBER:  PCT/DK97/00490
;  PRIOR FILING DATE:  1997-10-30
;  NUMBER OF SEQ ID NOS:  18
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 17
;   LENGTH: 664
;   TYPE: PRT
;   ORGANISM: Saccharomyces cerevisiae
US-09-295-186-17


  Query Match             15.8%;  Score 56.5;  DB 3;  Length 664;
  Best Local Similarity   28.4%;  Pred. No. 24;
  Matches   19;  Conservative   10;  Mismatches   19;  Indels   19;  Gaps    2;

Qy       9 VDGRSARGGCQTSND----------------AAAAALAASSAAAALQIIREKIPEKYLFQ 52
           :| ||  |    ||                 |:|:| |::||:|:     :|     |
Db     585 IDSRSVSG---VGNDDYSSSASLSASAAAASASASASASASASASGSSTHKKNAGNALVN 641

Qy      53 FHNLNSN 59
           : |||:|
Db     642 YSNLNTN 648


RESULT  13
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
;  GENERAL INFORMATION:
;    APPLICANT:  Leung, David W.
;    APPLICANT:  Weeks, Reitha
;    TITLE OF INVENTION:  MAMMALIAN
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cell Therapeutics, Inc.
;      STREET:  201 Elliott Avenue West
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98119
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
;      COMPUTER:  PC Clone (486 microprocessor)
;      OPERATING SYSTEM:  MS-DOS Version 6.1, Windows NT
;      SOFTWARE:  WORD 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,814D
;      FILING DATE:  28-Jun-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Faciszewski, Stephen
;      REGISTRATION NUMBER:  36,131
;      REFERENCE/DOCKET NUMBER:  1803
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)282-7100
;      TELEFAX:  (206)284-6206
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  446
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:  CDP-diacylglycerol synthase (CDS)
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-672-814D-11


  Query Match             15.6%;  Score 56;  DB 2;  Length 446;
  Best Local Similarity   34.0%;  Pred. No. 17;
  Matches   17;  Conservative    9;  Mismatches   22;  Indels    2;  Gaps    1;

Qy      20 TSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           : :|||    :|: ::  :    |||||:       |  || :  ||||
Db      20 SGSDAANKRNSAADSSDHVDSEEEKIPEEKFVD--ELAKNLPQGTDKTPE 67


RESULT  14
US-07-881-075-1
; Sequence 1, Application US/07881075
; Patent No. 5444149
;  GENERAL INFORMATION:
;    APPLICANT:  KEENE, JACK D.
;    APPLICANT:  KING, PETER H.
;    APPLICANT:  LEVINE, TODD
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS USEFUL IN THE
;    TITLE OF INVENTION:  RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
;    TITLE OF INVENTION:  INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
```

```
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/881,075
;      FILING DATE:  19920511
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5444149man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  714-154-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)521-4500
;      TELEFAX:  (703)486-2347
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  485 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-07-881-075-1


  Query Match             15.2%;  Score 54.5;  DB 1;  Length 485;
  Best Local Similarity   43.2%;  Pred. No. 30;
  Matches   16;  Conservative    6;  Mismatches   10;  Indels    5;  Gaps    1;

Qy       6 EFQVDGRSARGGCQT-----SNDAAAAALAASSAAAA 37
           :| :    | ||  |      : |||||:||::|||| 
Db       4 DFIMANTGAGGGVDTQAQLMQSAAAAAAVAATNAAAA 40


RESULT  15
US-08-120-827-1
; Sequence 1, Application US/08120827
; Patent No. 5525495
;  GENERAL INFORMATION:
;    APPLICANT:  KEENE, JACK D.
;    APPLICANT:  KING, PETER H.
;    APPLICANT:  LEVINE, TODD
;    TITLE OF INVENTION:  METHODS AND COMPOSITIONS USEFUL IN THE
;    TITLE OF INVENTION:  RECOGNITION, BINDING AND EXPRESSION OF RIBONUCLEIC ACIDS
;    TITLE OF INVENTION:  INVOLVED IN CELL GROWTH, NEOPLASIA AND IMMUNOREGULATION
;    NUMBER OF SEQUENCES:  101
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 Jefferson Davis Highway, Fourth Floor
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/120,827
;      FILING DATE:  15-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5525495man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  714-158-0 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)413-3000
;      TELEFAX:  (703)413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  485 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-120-827-1


  Query Match             15.2%;  Score 54.5;  DB 1;  Length 485;
  Best Local Similarity   43.2%;  Pred. No. 30;
  Matches   16;  Conservative    6;  Mismatches   10;  Indels    5;  Gaps    1;

Qy       6 EFQVDGRSARGGCQT-----SNDAAAAALAASSAAAA 37
           :| :    | ||  |      : |||||:||::|||| 
Db       4 DFIMANTGAGGGVDTQAQLMQSAAAAAAVAATNAAAA 40


Search completed: February  3, 2001, 02:19:42
Job time: 1704 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:26 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-3
Perfect score: 358
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	218	60.9	361	18	W34336	Tomato mottle viru
2	218	60.9	361	18	W34324	Tomato mottle viru
3	218	60.9	361	18	W34325	Tomato mottle viru
4	218	60.9	361	18	W34326	Tomato mottle viru
5	217	60.6	353	18	W34338	Bean golden mosaic
6	217	60.6	353	18	W34332	Bean golden mosaic
7	217	60.6	353	18	W34333	Bean golden mosaic
8	217	60.6	353	18	W34334	Bean golden mosaic
9	217	60.6	353	18	W34335	Bean golden mosaic
10	215	60.1	353	8	P70407	ORF 4 gene product
11	212	59.2	359	17	R88870	Sardinian tomato y
12	212	59.2	359	17	R88871	Sardinian tomato y
13	212	59.2	359	17	R88872	Sardinian tomato y
14	200.5	56.0	361	8	P70562	Product of ORF 4 f
15	198	55.3	362	19	W56495	Tobacco leaf curli
16	190	53.1	357	18	W34337	Tomato yellow leaf
17	190	53.1	357	18	W34329	Tomato yellow leaf
18	190	53.1	357	18	W34330	Tomato yellow leaf
19	190	53.1	357	18	W34331	Tomato yellow leaf
20	62.5	17.5	579	19	W63707	Human hSK2 protein
21	61.5	17.2	66	11	R08067	Synthetic antifree
22	61	17.0	1212	20	W88287	Rat (Na,K)-ATPase
23	61	17.0	1212	20	W88288	Rat (Na,K)-ATPase
24	61	17.0	1212	20	W88289	Rat (Na,K)-ATPase
25	61	17.0	1212	20	W88290	Rat (Na,K)-ATPase
26	60.5	16.9	131	18	W34327	Tomato mottle viru
27	60.5	16.9	509	17	R85311	Cystathionine gamm
28	60	16.8	314	18	W14283	Human neuroblastom
29	60	16.8	314	21	Y70572	Mouse Phox2b prote
30	59	16.5	3596	21	Y87407	Bordetella pertuss
31	59	16.5	3647	11	R05041	Filamentous haemag
32	58.5	16.3	512	19	W68473	HIV-1 strain YBF30
33	58.5	16.3	1713	16	R70148	Deduced sequence o
34	57	15.9	283	11	R03623	Zucchini yellow mo
35	57	15.9	3080	14	R35081	ZYMV polyprotein.
36	54.5	15.2	46	11	R08098	Synthetic antifree
37	54.5	15.2	483	14	R43893	elav. Drosophila
38	54.5	15.2	485	19	W60577	Drosophila neuron-
39	54.5	15.2	738	19	W56163	New DNA sequence i
40	54	15.1	52	20	W68001	Fragment of human
41	54	15.1	409	13	R25693	Heat resistant car
42	54	15.1	753	20	W93495	N. meningitidis st
43	53.5	14.9	41	11	R08066	Synthetic antifree
44	53.5	14.9	73	11	R08068	Synthetic antifree
45	53.5	14.9	264	20	Y14820	PanB amino acid se

ALIGNMENTS

RESULT 1
W34336
ID W34336 standard; Protein; 361 AA.
XX
AC W34336;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 protein.
XX
KW Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93294.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

XX
PS Example 3.3; Page 57-58; 132pp; English.
XX
CC This protein comprises the wild-type AC1 protein of tomato mottle
CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1
CC gene (see T93294) must be expressed for efficient replication of
CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a
CC DNA binding site specific to the DNA-A common region, a DNA nicking
CC activity, and an NTP binding activity. The invention involves
CC production of transgenic plants containing DNA comprising AC1 or C1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC W34324-35) that have mutations in the highly conserved DNA-nicking
CC domain and/or the NTP-binding domain.
XX
SQ Sequence 361 AA;

Query Match 60.9%; Score 218; DB 18; Length 361;
Best Local Similarity 57.1%; Pred. No. 2e-20;
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
|: ||:||:|||||||| |::||: | || ||| :|| ::||: |: :: | ||: |||
Db 110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy 61 DRIFDKTPEP 70
:||| | |||
Db 170 erifakapep 179

RESULT 2
W34324
ID W34324 standard; Protein; 361 AA.
XX
AC W34324;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93282.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 3.4; Page 60-62; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant, designated
CC ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC W34336). It is encoded by a mutant AC1 gene (see T93282) of
CC tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC NTP-binding domains. The AC1 gene (see also T93294) must be
CC expressed for efficient replication of the two genomic components,
CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC involves production of transgenic plants containing DNA comprising
CC geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC interfere in trans with geminiviral replication during infection.
CC Such transgenic plants are resistant to viral infection. The
CC AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC polypeptides (see W34324-35) that have mutations in the highly
CC conserved DNA-nicking and/or NTP-binding domains.
XX
SQ Sequence 361 AA;

Query Match 60.9%; Score 218; DB 18; Length 361;
Best Local Similarity 57.1%; Pred. No. 2e-20;
Matches 40; Conservative 15; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
|: ||:||:|||||||| |::||: | || ||| :|| ::||: |: :: | ||: |||
Db 110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy 61 DRIFDKTPEP 70
:||| | |||
Db 170 erifakapep 179

RESULT 3
W34325
ID W34325 standard; Protein; 361 AA.
XX
AC W34325;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93283.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 3.5; Page 64-65; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant, designated
CC ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see

```
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.9%;  Score 218;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 2e-20;
  Matches   40;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:||||||||| |::||: | || |||  :|| ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT   4
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.9%;  Score 218;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 2e-20;
  Matches   40;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:||||||||| |::||: | || |||  :|| ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT   5
W34338
ID   W34338 standard; Protein; 353 AA.
XX
AC   W34338;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 protein.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93314.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
```

```
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.6%;  Score 217;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-20;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179


RESULT   6
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.6%;  Score 217;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-20;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179


RESULT   7
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.6%;  Score 217;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-20;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
```

```
          |: ||:||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179
```

RESULT 8
W34334
ID W34334 standard; Protein; 353 AA.
XX
AC W34334;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93292.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 111-112; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminiviral
CC replication during infection. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGMV,
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the
CC highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ Sequence 353 AA;

```
  Query Match             60.6%;  Score 217;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-20;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179
```

RESULT 9
W34335
ID W34335 standard; Protein; 353 AA.
XX
AC W34335;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93293.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 115-116; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminiviral
CC replication during infection. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGMV,
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the
CC highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ Sequence 353 AA;

```
  Query Match             60.6%;  Score 217;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-20;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179
```

RESULT 10
P70407
ID P70407 standard; Protein; 353 AA.
XX

```
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             60.1%;  Score 215;  DB 8;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 4.8e-20;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:|||||||||||| |::||: | || | |  :|| |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  11
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             59.2%;  Score 212;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-19;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT  12
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
```

```
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             59.2%;  Score 212;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-19;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT  13
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             59.2%;  Score 212;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.2e-19;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| | |: | | : || :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             56.0%;  Score 200.5;  DB 8;  Length 361;
  Best Local Similarity   63.1%;  Pred. No. 3.9e-18;
  Matches   41;  Conservative    8;  Mismatches   15;  Indels    1;  Gaps    1;

Qy      5 GEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLDRIF 64
          | |||||||||||  |::||| | || : |   || |:||| |::|: |||||| || |||
Db    114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy     65 -DKTP 68
           |: |
Db    174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             55.3%;  Score 198;  DB 19;  Length 362;
  Best Local Similarity   50.0%;  Pred. No. 8.2e-18;
  Matches   42;  Conservative   12;  Mismatches   14;  Indels   16;  Gaps    1;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAAALAASSAAAALQIIREKIPEKYLFQFHNLNSNLD 61
          | :| |||||||||  ||::||| | |: : | ::|| |:||| |: :: ||||||||||
Db    111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy     62 RI----------------FDKTPE 69
          ||                 ||: ||
Db    171 rifapplevfvcpfssssfdqvpe 194



Search completed: February  3, 2001, 02:15:27
Job time: 3434 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:50 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	281	77.8	226	12	O09727	O09727 leonurus mo
2	281	77.8	226	12	Q9WHF6	Q9whf6 tomato mild
3	279	77.3	361	12	Q67574	Q67574 bean golden
4	277	76.7	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	274	75.9	364	12	Q9QS55	Q9qs55 sweet potat
6	271	75.1	185	12	Q98693	Q98693 sida golden
7	266	73.7	149	12	P88975	P88975 macroptiliu
8	266	73.7	233	12	Q9YLA4	Q9yla4 macroptiliu
9	260	72.0	234	12	O39180	O39180 potato yell
10	254	70.4	190	12	Q9Z089	Q9z089 tobacco lea
11	254	70.4	190	12	Q9Z084	Q9z084 tobacco lea
12	253	70.1	190	12	Q9W827	Q9w827 tobacco lea
13	253	70.1	208	12	Q9Z0C4	Q9z0c4 tobacco lea
14	251	69.5	208	12	Q9Z0C0	Q9z0c0 tobacco lea
15	251	69.5	208	12	Q9Z0B8	Q9z0b8 tobacco lea
16	251	69.5	363	12	O72705	O72705 cotton leaf
17	251	69.5	363	12	O72719	O72719 cotton leaf
18	249	69.0	208	12	Q9Z0B6	Q9z0b6 tobacco lea
19	248	68.7	203	12	Q9Z083	Q9z083 tobacco lea
20	248	68.7	359	12	Q9YZV4	Q9yzv4 tomato yell
21	248	68.7	359	12	Q9YZV2	Q9yzv2 tomato yell
22	248	68.7	359	12	Q9YUX7	Q9yux7 tomato yell
23	248	68.7	359	12	Q9YL27	Q9yl27 tomato yell
24	248	68.7	363	12	O73577	O73577 cotton leaf
25	247	68.4	359	12	Q88942	Q88942 tomato yell
26	245	67.9	307	12	Q9IET7	Q9iet7 cotton leaf
27	245	67.9	361	12	O72723	O72723 cotton leaf
28	243	67.3	362	12	Q9IN48	Q9in48 okra enatio
29	242	67.0	349	12	Q88888	Q88888 tomato pseu
30	242	67.0	362	12	O56816	O56816 chayote mos
31	241	66.8	208	12	Q9Z0A0	Q9z0a0 tobacco lea
32	241	66.8	354	12	Q9IN42	Q9in42 south afric
33	240	66.5	190	12	Q9Z086	Q9z086 tobacco lea
34	240	66.5	208	12	Q9Z0C6	Q9z0c6 tobacco lea
35	240	66.5	358	12	Q9JEA2	Q9jea2 cassava gem
36	239	66.2	363	12	O73494	O73494 okra yellow
37	237	65.7	306	12	Q9IET1	Q9iet1 cotton leaf
38	237	65.7	353	12	O72692	O72692 beet curly
39	237	65.7	354	12	Q65438	Q65438 beet curly
40	237	65.7	360	12	Q9QCV3	Q9qcv3 tomato leaf
41	236	65.4	190	12	Q9Z0A7	Q9z0a7 tobacco lea
42	236	65.4	360	12	O72707	O72707 cotton leaf
43	236	65.4	360	12	O72710	O72710 cotton leaf
44	236	65.4	360	12	Q65317	Q65317 ageratum ye
45	236	65.4	362	12	Q9YYT7	Q9yyt7 althea rose

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.8%; Score 281; DB 12; Length 226;
Best Local Similarity 78.6%; Pred. No. 2.9e-25;
Matches 55; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: |||||||||||:||| || ||||||||||  |  |||||   :||||||||||||||
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             77.8%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   74.3%;  Pred. No. 2.9e-25;
  Matches   52;  Conservative   11;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:|||||||||||||||::||   :|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.3%;  Score 279;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 8.5e-25;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           || |||||||||| ||:||||:||||||||||:|||   :|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.7%;  Score 277;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 8.4e-25;
  Matches   52;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| || ||||:||||||||||:|||   :|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | ||.|
Db     173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.9%;  Score 274;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 3.3e-24;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||||||||||||||| ||:|||||||||| ||| ||||   :|||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.1%;  Score 271;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 3.4e-24;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:||   :||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.7%;  Score 266;  DB 12;  Length 149;
  Best Local Similarity   70.0%;  Pred. No. 1e-23;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|:   :|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.7%;  Score 266;  DB 12;  Length 233;
  Best Local Similarity   70.0%;  Pred. No. 1.7e-23;
```

```
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||||||||||||: :|| |::|:   :|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   9
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234     234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             72.0%;  Score 260;  DB 12;  Length 234;
  Best Local Similarity   68.6%;  Pred. No. 8.7e-23;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| ||:|||||||||: :||||::||   :|||:|||:|||:|:|
Db    110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy     61 DRIFDKTPEP 70
          |||| | | |
Db    170 DRIFMKAPNP 179


RESULT  10
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1       1
FT   NON_TER     190 .   190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             70.4%;  Score 254;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 3.5e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          || || ||:|||||||||| :||| ||||||||| ||| ||   :|: |::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1       1
FT   NON_TER     190     190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             70.4%;  Score 254;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 3.5e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          || || ||:|||||||||| :||| ||||||||| ||| ||   :|: |::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             70.1%;  Score 253;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 4.5e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| ||||||:|| ||| ||   :|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  13
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             70.1%;  Score 253;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 5e-22;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || |||||||||||||| :||| ||||||:|| ||| ||   :|: ::||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  14
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             69.5%;  Score 251;  DB 12;  Length 208;
  Best Local Similarity   57.6%;  Pred. No. 8.7e-22;
  Matches   49;  Conservative    9;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||| :|| ||   :|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;
```

Query Match 69.5%; Score 251; DB 12; Length 208;
Best Local Similarity 57.6%; Pred. No. 8.7e-22;
Matches 49; Conservative 9; Mismatches 11; Indels 16; Gaps 1;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||| ||||||||| :|| ||   :|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186
```

Search completed: February 3, 2001, 02:22:51
Job time: 1888 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:25 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	346	95.8	352	1	VAL1_TGMV	P03567 tomato gold
2	249	69.0	361	1	VAL1_PYMVV	P27258 potato yell
3	240	66.5	358	1	VAL1_CLVK	P14982 cassava lat
4	240	66.5	358	1	VAL1_CLVN	P14972 cassava lat
5	233	64.5	362	1	VAL1_TYLCA	P36279 tomato yell
6	229	63.4	359	1	VAL1_TYLCU	P38609 tomato yell
7	222	61.5	349	1	VAL1_PHUV	Q06923 pepper huas
8	222	61.5	359	1	VAL1_TYLCM	P27260 tomato yell
9	217	60.1	353	1	VAL1_BGMV	P05175 bean golden
10	216	59.8	355	1	VAL1_ABMVW	P21947 abutilon mo
11	213	59.0	358	1	VAL1_BCTV	P14991 beet curly
12	210	58.2	361	1	VAL1_TMOV	Q06657 tomato mott
13	200	55.4	357	1	VAL1_TYLCV	P27259 tomato yell
14	128	35.5	347	1	VAL1_SLCV	P29048 squash leaf
15	61.5	17.0	299	1	Y175_HELPJ	Q9zmq7 helicobacte
16	61.5	17.0	334	1	G3P_BACST	P00362 bacillus st
17	61.5	17.0	335	1	G3P_BACCO	P15115 bacillus co
18	59	16.3	1483	1	CYP1_YEAST	P12351 saccharomyc
19	58.5	16.2	297	1	RRPP_RABVP	P06747 rabies viru
20	58	16.1	555	1	GLPD_BACSU	P18158 bacillus su
21	57	15.8	223	1	DGK2_LACAC	Q59484 lactobacill
22	57	15.8	316	1	SPEE_COFAR	O82147 coffea arab
23	57	15.8	379	1	CYB_ASTPE	Q33818 asterina pe
24	57	15.8	387	1	Y4PF_RHISN	P55615 rhizobium s
25	57	15.8	836	1	GCSR_HUMAN	Q99062 homo sapien
26	57	15.8	1287	1	SKI2_YEAST	P35207 saccharomyc
27	56	15.5	367	1	LHX4_MOUSE	P53776 mus musculu
28	56	15.5	577	1	PTLB_LACCA	P24400 lactobacill
29	56	15.5	874	1	SLAP_BACLI	P49052 bacillus li
30	55.5	15.4	299	1	Y175_HELPY	P56112 helicobacte
31	55.5	15.4	323	1	VAL1_PASVK	Q00338 panicum str
32	55.5	15.4	397	1	METL_RAT	P13444 rattus norv
33	55.5	15.4	428	1	GBA1_CANAL	P28868 candida alb
34	55	15.2	355	1	CRTB_RHOSH	P54905 rhodobacter
35	55	15.2	524	1	MPPA_RAT	P20069 rattus norv
36	55	15.2	525	1	MPPA_HUMAN	Q10713 homo sapien
37	54.5	15.1	297	1	RRPP_RABVE	P22559 rabies viru
38	54.5	15.1	405	1	WCAC_ECOLI	P71237 escherichia
39	54.5	15.1	445	1	ACSC_MOOTH	Q07340 moorella th
40	54.5	15.1	1108	1	KBQ9_YEAST	P38080 saccharomyc
41	54.5	15.1	1232	1	YOQ5_CAEEL	P34643 caenorhabdi
42	54.5	15.1	1713	1	LMA3_HUMAN	Q16787 homo sapien
43	54	15.0	111	1	RL22_ACHLA	P29222 acholeplasm
44	54	15.0	258	1	YL48_MYCTU	O06228 mycobacteri
45	54	15.0	274	1	YK53_DEIRA	Q9rss1 deinococcus

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.8%; Score 346; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.9e-34;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             69.0%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 1.1e-22;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: || ||:||||||| || |||||||||: :|| |::||   :|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             66.5%;  Score 240;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.3e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :|   :|:|:: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             66.5%;  Score 240;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.3e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :|   :|:|:: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             64.5%;  Score 233;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 9.2e-21;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           || ||||:|||||||| |::||| |:|||  || ||| ::    |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             63.4%;  Score 229;  DB 1;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 2.8e-20;
  Matches   42;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|    |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created).
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             61.5%;  Score 222;  DB 1;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.9e-19;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||||:|||||||| |::|| |:|||::| |||||||    |: :  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             61.5%;  Score 222;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.9e-19;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV     STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.1%;  Score 217;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 7.7e-19;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:|||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db    110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             59.8%;  Score 216;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 1e-18;
  Matches   41;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: |:|||| : || |:  |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.0%;  Score 213;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 2.4e-18;
  Matches   39;  Conservative   14;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:|||| :| ::||||:    |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             58.2%;  Score 210;  DB 1;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 5.5e-18;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::    |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             55.4%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 8.8e-17;
  Matches   39;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           :|  |:||||||||| |::||| |||||: || ||| |:    |: |: |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             35.5%;  Score 128;  DB 1;  Length 347;
  Best Local Similarity   39.4%;  Pred. No. 4.3e-08;
  Matches   26;  Conservative   13;  Mismatches   23;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| || |  |: ::  |:||| :|:::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
Y175_HELPJ
ID   Y175_HELPJ     STANDARD;      PRT;   299 AA.
AC   Q9ZMQ7;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN HP0175 PRECURSOR.
GN   HP0175 OR JHP0161.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC       STRONG, TO C.JEJUNI CBF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001454; AAD05744.1; -.
DR   HSSP; Q13526; 1PIN.
DR   INTERPRO; IPR000297; -.
DR   PFAM; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
KW   Hypothetical protein; Isomerase; Rotamase; Signal.
FT   SIGNAL       1     21       POTENTIAL.
FT   CHAIN       22    299       HYPOTHETICAL PROTEIN HP0175.
FT   DOMAIN     154    253       PPIC-LIKE.
SQ   SEQUENCE   299 AA;  34040 MW;  9C037B1CD1110143 CRC64;


  Query Match             17.0%;  Score 61.5;  DB 1;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 4;
  Matches   18;  Conservative   12;  Mismatches   13;  Indels   21;  Gaps    2;
```

```
Qy      22 NDAAAEALNAS------------------SKEEALQIIAAAIPEKYLFQFHNLNSNLDR 62
           |:| || || :                   :|::| ::    ||||  :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154


Search completed: February  3, 2001, 02:24:26
Job time: 638 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:38 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	346	95.8	352	1	QQCVL1	AL1 protein - toma
2	249	69.0	361	1	QQCVPT	AL1 protein - pota
3	240	66.5	358	2	S07594	hypothetical prote
4	233	64.5	362	1	JQ1887	AL1 protein - toma
5	229	63.4	359	2	S39211	gene C1 protein -
6	222	61.5	349	2	JQ2300	replicase - pepper
7	222	61.5	349	2	S31875	AL1 protein - pepp
8	222	61.5	359	2	S22593	hypothetical prote
9	219	60.7	359	2	S39235	gene C1 protein -
10	216	59.8	351	2	JQ2327	AL1 protein - Indi
11	216	59.8	355	1	QQCVW1	AV1 protein - abut
12	213	59.0	385	2	S28360	AL1 protein - beet
13	210	58.2	358	1	JQ1870	AL1 protein - toma
14	200	55.4	357	1	QQCVC1	AL1 protein - toma
15	199	55.1	360	2	S59885	replication-associ
16	134	37.1	131	2	S45059	AC1 protein (clone
17	128	35.5	347	1	QQCVS1	AL1 protein - squa
18	72	19.9	587	2	JC1419	Fc gamma (IgG) rec
19	63.5	17.6	584	2	T19061	hypothetical prote
20	62.5	17.3	1006	2	S74992	hypothetical prote
21	61.5	17.0	299	2	B71967	probable peptidyl-
22	61.5	17.0	335	2	JS0164	glyceraldehyde-3-p
23	61.5	17.0	481	2	A70091	probable phosphoes
24	61.5	17.0	1044	2	T43800	protein kinase bub
25	61.5	17.0	1203	2	T17415	mycelial surface a
26	61.5	17.0	1229	2	T48959	kinesin-like prote
27	60.5	16.8	159	2	C71838	NADH dehydrogenase
28	60.5	16.8	159	2	E64677	NADH dehydrogenase
29	60.5	16.8	334	1	DEBSGF	glyceraldehyde-3-p
30	60	16.6	767	2	F71479	hypothetical prote
31	60	16.6	1008	2	T41244	SEC14 protein homo
32	59	16.3	160	2	G82060	hypothetical prote
33	59	16.3	1502	1	RGBYH1	CYC1/CYP3 transcri
34	58.5	16.2	297	1	MNVNRV	nonstructural prot
35	58.5	16.2	357	2	T02246	hypothetical prote
36	58.5	16.2	769	2	F81742	conserved hypothet
37	58.5	16.2	2135	2	T14602	variant-specific s
38	58	16.1	555	2	C45868	glycerol-3-phospha
39	57.5	15.9	858	2	C64577	virB4 homolog - He
40	57.5	15.9	1409	2	S74916	alkaline phosphata
41	57	15.8	224	2	B56168	deoxyguanosine kin
42	57	15.8	380	2	S70594	ubiquinol--cytochr
43	57	15.8	604	2	T19682	hypothetical prote
44	57	15.8	771	2	B38252	granulocyte colony
45	57	15.8	783	2	JH0329	granulocyte colony

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.8%; Score 346; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 2.6e-33;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             69.0%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 8.3e-22;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: || ||:||||||||| || ||||||||| :|| |::||   :|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             66.5%;  Score 240;  DB 2;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 9.5e-21;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:||: || ||| :|   :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.5%;  Score 233;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 6.5e-20;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          || |||||:|||||||| |::||| |:|||  || ||| ::    |: |: |||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.4%;  Score 229;  DB 2;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 1.9e-19;
  Matches   42;  Conservative    8;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|    |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 222;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-18;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| |::||  |:|||::| |||||||    |: :  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 222;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-18;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||    |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.5%;  Score 222;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.3e-18;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    | |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 219;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.9e-18;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|    | |:  |||::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  10
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 216;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 6.5e-18;
  Matches   41;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| | |||: || |||:|:    |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  11
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 216;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 6.6e-18;
  Matches   41;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
```

```
          |  |||||:||||||||| ||:||: |:||||   : || |:    |: |: | ||: |||
Db    110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179


RESULT  12
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.0%;  Score 213;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.6e-17;
  Matches   39;  Conservative   14;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| ||:||: |:||||:| ::||||:    |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  13
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 210;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 3.4e-17;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||| | : || ::    |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176


RESULT  14
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             55.4%;  Score 200;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 5.2e-16;
  Matches   39;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
          :|  |:|||||||| |::||| |||||: || ||| |:    |: |:|||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


RESULT  15
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             55.1%;  Score 199;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 6.9e-16;
  Matches   38;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
          :| ||:|||||||| |::||| |||:|: ||  || |:    |: ::|||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178


Search completed: February  3, 2001, 02:17:38
Job time: 2210 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:42 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	222	61.5	359	3	US-08-809-103B-2	Sequence 2, Appli
2	222	61.5	359	3	US-08-809-103B-4	Sequence 4, Appli
3	222	61.5	359	3	US-08-809-103B-6	Sequence 6, Appli
4	222	61.5	359	3	US-08-809-103B-8	Sequence 8, Appli
5	60.5	16.8	334	5	5290690-11	Patent No. 5290690
6	58.5	16.2	356	1	US-08-781-562-1	Sequence 1, Appli
7	57	15.8	602	2	US-08-419-652-6	Sequence 6, Appli
8	57	15.8	771	1	US-07-923-976-6	Sequence 6, Appli
9	57	15.8	783	5	5422248-2	Patent No. 5422248
10	57	15.8	836	1	US-07-923-976-4	Sequence 4, Appli
11	57	15.8	863	1	US-07-923-976-8	Sequence 8, Appli
12	55	15.2	783	5	5231168-2	Patent No. 5231168
13	54.5	15.1	1713	3	US-08-600-982-24	Sequence 24, Appl
14	54.5	15.1	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
15	54	15.0	271	1	US-08-276-919-10	Sequence 10, Appl
16	54	15.0	271	1	US-08-776-088-13	Sequence 13, Appl
17	54	15.0	271	4	PCT-US95-09145A-13	Sequence 13, Appl
18	54	15.0	274	1	US-08-776-088-15	Sequence 15, Appl
19	54	15.0	274	4	PCT-US95-09145A-15	Sequence 15, Appl
20	54	15.0	325	1	US-08-276-919-4	Sequence 4, Appli
21	54	15.0	325	1	US-08-276-919-13	Sequence 13, Appl
22	54	15.0	325	1	US-08-776-088-4	Sequence 4, Appli
23	54	15.0	325	1	US-08-776-088-18	Sequence 18, Appl
24	54	15.0	325	4	PCT-US95-09145A-4	Sequence 4, Appli
25	54	15.0	325	4	PCT-US95-09145A-18	Sequence 18, Appl
26	54	15.0	354	1	US-08-276-919-2	Sequence 2, Appli
27	54	15.0	354	1	US-08-776-088-2	Sequence 2, Appli
28	54	15.0	354	1	US-08-776-088-6	Sequence 6, Appli
29	54	15.0	354	3	US-09-325-320-2	Sequence 2, Appli
30	54	15.0	354	4	PCT-US95-09145A-2	Sequence 2, Appli
31	54	15.0	354	4	PCT-US95-09145A-6	Sequence 6, Appli
32	53	14.7	297	2	US-09-027-013-3	Sequence 3, Appli
33	53	14.7	297	3	US-09-244-233-3	Sequence 3, Appli
34	53	14.7	844	3	US-09-029-267-20	Sequence 20, Appl
35	53	14.7	852	1	US-08-190-802A-59	Sequence 59, Appl
36	53	14.7	852	1	US-08-190-802A-67	Sequence 67, Appl
37	52.5	14.5	3898	2	US-08-876-991-2	Sequence 2, Appli
38	52.5	14.5	3898	2	US-09-059-853-2	Sequence 2, Appli
39	52	14.4	207	2	US-08-177-109A-60	Sequence 60, Appl
40	52	14.4	207	2	US-08-687-706-60	Sequence 60, Appl
41	52	14.4	454	3	US-08-446-100-31	Sequence 31, Appl
42	52	14.4	482	1	US-07-792-885A-1	Sequence 1, Appli
43	52	14.4	482	1	US-08-142-439A-7	Sequence 7, Appli
44	52	14.4	482	2	US-08-869-477-7	Sequence 7, Appli
45	51.5	14.3	870	2	US-08-266-311-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
; GENERAL INFORMATION:
;   APPLICANT: GRONENBORN, Bruno
;   TITLE OF INVENTION: PHYTOPATHOGENIC DNA VIRUS RESISTANT
;   TITLE OF INVENTION: TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;   NUMBER OF SEQUENCES: 17
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: YOUNG & THOMPSON
;     STREET: 745 South 23rd Street
;     CITY: Arlington
;     STATE: Virginia
;     COUNTRY: U.S.A.
;     ZIP: 22202
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/809,103B
;     FILING DATE: 17-MAR-1997
;     CLASSIFICATION: 800
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: FR 94.11040
;     FILING DATE: 15-SEP-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: WO PCT/FR95/01192
;     FILING DATE: 15-SEP-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: PATCH, Andrew J.
;     REGISTRATION NUMBER: 32,925
;     REFERENCE/DOCKET NUMBER: USB94AL CNR TOM
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (703) 521-2297
;     TELEFAX: (703) 685-0573
;     TELEX: 248425 EMBON
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 359 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-809-103B-2


  Query Match 61.5%; Score 222; DB 3; Length 359;
  Best Local Similarity 58.0%; Pred. No. 2.6e-22;
  Matches 40; Conservative 10; Mismatches 19; Indels 0; Gaps 0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             61.5%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.6e-22;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             61.5%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.6e-22;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             61.5%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.6e-22;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
5290690-11
;Patent No. 5290690
;     APPLICANT: MRABET, NADIR;LASTERS, IGNACE;STANSSENS, PATRICK
;MATTHYSSENS, GASTON;WODAK, SHOSHANA;QUAX, WILHELMUS J.
;     TITLE OF INVENTION: METHODS AND MEANS FOR CONTROLLING THE
;STABILITY OF PROTEINS
;     NUMBER OF SEQUENCES: 22
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/398,706
;       FILING DATE: 25-AUG-1989
;SEQ ID NO:11:
;       LENGTH: 334
5290690-11


  Query Match             16.8%;  Score 60.5;  DB 5;  Length 334;
  Best Local Similarity   28.4%;  Pred. No. 2.6;
  Matches   19;  Conservative   11;  Mismatches   28;  Indels    9;  Gaps    1;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKE---------EALQIIAAAIPEKYLFQ 52
           | |||  ||       |  |  : ||: | | :|:         | : ::     :||  :
Db      82 LAWGEIGVDIVVESTGRFTKREDAAKHLEAGAKKVIISAPAKVENITVVMGVNQDKYDPK 141

Qy      53 FHNLNSN 59
             |:: ||
Db     142 AHHVISN 148


RESULT   6
US-08-781-562-1
; Sequence 1, Application US/08781562
; Patent No. 5763589
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN MEMBRANE PROTEIN
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/781,562
;      FILING DATE:  Herewith
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0181 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  356 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  Consensus
;      CLONE:  Consensus
US-08-781-562-1


  Query Match             16.2%;  Score 58.5;  DB 1;  Length 356;
  Best Local Similarity   27.6%;  Pred. No. 5.3;
  Matches   16;  Conservative   12;  Mismatches   17;  Indels   13;  Gaps    1;

Qy      19 QTSNDAAAEALNASSKEEALQIIAAA----------IPEKYLFQFHNLNSNLDRI 63
           | : :|:|    | :| ||::|:|||           : |:|:  |  |  : : |
Db     239 QAAGEASAVLAKAKAKAEAIRILAAALTQHNGDAAASLTVAEQYVSAFSKLAKDSNTI 296


RESULT   7
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
;  GENERAL INFORMATION:
;    APPLICANT:  Chua, Anne O
;    APPLICANT:  Gubler, Ulrich A
;    TITLE OF INVENTION:  INTERLEUKIN-12 RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
```

```
;       STATE:  New Jersey
;       COUNTRY:  United States of America
;       ZIP:  07110-1199
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  PC compatible
;       OPERATING SYSTEM:  MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/419,652
;       FILING DATE:  11-APR-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/248,532
;       FILING DATE:  31-MAY-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/094,713
;       FILING DATE:  19-JUL-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Kass, Alan P
;       REGISTRATION NUMBER:  32142
;       REFERENCE/DOCKET NUMBER:  CD 9174
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (201) 235-4205
;       TELEFAX:  (201) 235-3500
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  602 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
;     FEATURE:
;       NAME/KEY:  Region
;       LOCATION:  1..602
;       OTHER INFORMATION:  /note= "Represents residues 98 to
;       OTHER INFORMATION:  731 of human granulocyte colony-stimulating
;       OTHER INFORMATION:  factor-receptor."
US-08-419-652-6


  Query Match             15.8%;  Score 57;  DB 2;  Length 602;
  Best Local Similarity   24.0%;  Pred. No. 17;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNL 56
           |  :    :|| |||  |:  : :   :          || |:|  : |:
Db      37 FTLKSFKSRGNCQTQGDSILDCVPKDGQSH------CCIPRKHLLLYQNM 80


RESULT   8
US-07-923-976-6
; Sequence 6, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-6


  Query Match             15.8%;  Score 57;  DB 1;  Length 771;
  Best Local Similarity   24.0%;  Pred. No. 24;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNL 56
           |  :    :|| |||  |:  : :   :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSH------CCIPRKHLLLYQNM 199


RESULT   9
5422248-2
;Patent No. 5422248
;    APPLICANT: SMITH, CRAIG A.;LARSEN, ALF D.;SIMS, JOHN E.;
;BENSON, CURTIS M.
;    TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
;STIMULATING FACTOR RECEPTORS
;    NUMBER OF SEQUENCES: 6
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/6,183
;      FILING DATE: 15-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 587,329
;      FILING DATE: 24-SEP-1990
;      APPLICATION NUMBER: 522,952
;      FILING DATE: 03-APR-1990
;      APPLICATION NUMBER: 416,306
;      FILING DATE: 03-OCT-1989
;      APPLICATION NUMBER: 412,816
;      FILING DATE: 26-SEP-1989
;SEQ ID NO:2:
;      LENGTH: 783
5422248-2


  Query Match             15.8%;  Score 57;  DB 5;  Length 783;
  Best Local Similarity   24.0%;  Pred. No. 24;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNL 56
           |  :    :|| |||  |:  : :   :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSH------CCIPRKHLLLYQNM 199


RESULT  10
US-07-923-976-4
; Sequence 4, Application US/07923976
; Patent No. 5574136
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  836 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-4


  Query Match             15.8%;  Score 57;  DB 1;  Length 836;
  Best Local Similarity   24.0%;  Pred. No. 27;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNL 56
           | :    :|| |||  |:  : :    :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSH------CCIPRKHLLLYQNM 199


RESULT  11
US-07-923-976-8
; Sequence 8, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  863 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-8


  Query Match             15.8%;  Score 57;  DB 1;  Length 863;
  Best Local Similarity   24.0%;  Pred. No. 28;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNL 56
           | :    :|| |||  |:  : :    :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSH------CCIPRKHLLLYQNM 199


RESULT  12
5231168-2
;Patent No. 5231168
;    APPLICANT: DZIEGIEL, MORTEN;BORRE, MARTIN;JEPSEN, SOREN;
;VUUST, JENS;RIENECK, KLAUS;WIND, ANNETTE;JAKOBSEN, PALLE H.
;    TITLE OF INVENTION: MALARIA ANTIGEN
;    NUMBER OF SEQUENCES: 19
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/409,658
;      FILING DATE: 18-SEP-1989
;SEQ ID NO:2:
;      LENGTH: 783
5231168-2


  Query Match             15.2%;  Score 55;  DB 5;  Length 783;
  Best Local Similarity   26.9%;  Pred. No. 46;
  Matches   14;  Conservative   10;  Mismatches   22;  Indels    6;  Gaps    1;

Qy      19 QTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRIFDKTPEP 70
           ||:|:   | :    |:    :   |||:: :    | | |:     ||:|
Db     614 QTNNNENIETIKPEEKKNEFSVEEKAIPQEPVVPTLNENENV------TPKP 659


RESULT  13
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
```

```
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             15.1%;  Score 54.5;  DB 3;  Length 1713;
  Best Local Similarity   32.8%;  Pred. No. 1.5e+02;
  Matches   20;  Conservative   11;  Mismatches   23;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIAAAIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|   :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  14
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24


  Query Match             15.1%;  Score 54.5;  DB 4;  Length 1713;
  Best Local Similarity   32.8%;  Pred. No. 1.5e+02;
  Matches   20;  Conservative   11;  Mismatches   23;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIAAAIPEKYLFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|   :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT  15
US-08-276-919-10
; Sequence 10, Application US/08276919
; Patent No. 5589579
;  GENERAL INFORMATION:
;    APPLICANT:  Torczynski, Richard M.
;    APPLICANT:  Bollon, Arthur P.
;    TITLE OF INVENTION:  Lung Cancer Marker
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  RICHARDS, MEDLOCK & ANDREWS
;      STREET:  1201 Elm Street, Suite 4500
;      CITY:  Dallas
;      STATE:  TX
;      COUNTRY:  US
;      ZIP:  75270-2197
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/276,919
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nelligan, Mark C.
;      REGISTRATION NUMBER:  36,389
;      REFERENCE/DOCKET NUMBER:  B35792
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  214-939-4500
;      TELEFAX:  214-939-4600
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  271 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-276-919-10
```

Query Match 15.0%; Score 54; DB 1; Length 271;
Best Local Similarity 29.5%; Pred. No. 15;
Matches 13; Conservative 10; Mismatches 15; Indels 6; Gaps 1;

```
Qy      21 SNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRIF 64
           ::|   :|  ||:| | | ::|       | :  : | : |:||
Db     121 NSDLYPDASTASNKSEGLAVLAV------LIEMGSFNPSYDKIF 158
```

Search completed: February 3, 2001, 02:19:43
Job time: 1705 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:27 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-4
Perfect score: 361
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	222	61.5	359	17	R88870	Sardinian tomato y
2	222	61.5	359	17	R88871	Sardinian tomato y
3	222	61.5	359	17	R88872	Sardinian tomato y
4	219	60.7	353	18	W34338	Bean golden mosaic
5	219	60.7	353	18	W34332	Bean golden mosaic
6	219	60.7	353	18	W34333	Bean golden mosaic
7	219	60.7	353	18	W34334	Bean golden mosaic
8	219	60.7	353	18	W34335	Bean golden mosaic
9	217	60.1	353	8	P70407	ORF 4 gene product
10	214	59.3	361	18	W34336	Tomato mottle viru
11	214	59.3	361	18	W34324	Tomato mottle viru
12	214	59.3	361	18	W34325	Tomato mottle viru
13	214	59.3	361	18	W34326	Tomato mottle viru
14	203.5	56.4	361	8	P70562	Product of ORF 4 f
15	200	55.4	357	18	W34329	Tomato yellow leaf
16	200	55.4	357	18	W34330	Tomato yellow leaf
17	200	55.4	357	18	W34331	Tomato yellow leaf
18	198	54.8	362	19	W56495	Tobacco leaf curli
19	192	53.2	357	18	W34337	Tomato yellow leaf
20	65.5	18.1	512	19	W68473	HIV-1 strain YBF30
21	61.5	17.0	299	18	W55450	H. pylori ORF 02ae
22	61.5	17.0	1203	21	Y83275	Candida albicans C
23	61.5	17.0	1203	21	Y84814	Amino acid sequenc
24	60.5	16.8	131	18	W34327	Tomato mottle viru
25	60.5	16.8	334	11	R03211	Amino acid sequenc
26	60	16.6	214	20	Y37449	Amino acid sequenc
27	58.5	16.2	297	8	P70318	Protein M1 encoded
28	58.5	16.2	306	20	W78284	Fragment of human
29	58.5	16.2	356	19	W57232	Human integral mem
30	58.5	16.2	356	20	W78160	Human secreted pro
31	58	16.1	108	16	R82847	Murine mutant cyto
32	57.5	15.9	204	20	W89814	Protein encoded by
33	57.5	15.9	299	20	W89829	Protein encoded by
34	57.5	15.9	299	20	W89849	Antigen from clust
35	57.5	15.9	304	19	W98862	H. pylori GHPO 167
36	57	15.8	108	16	R82846	Murine mutant cyto
37	57	15.8	213	17	W00648	Human G-CSF recept
38	57	15.8	333	17	R99141	Ligand binding cyt
39	57	15.8	603	21	Y32222	Granulocyte colony
40	57	15.8	771	12	R14256	Human GCSF recepto
41	57	15.8	783	12	R11741	Granulocyte colony
42	57	15.8	783	16	R77868	Human placental G-
43	57	15.8	783	18	W10485	Human granulocyte
44	57	15.8	800	18	W10486	Human granulocyte
45	57	15.8	801	12	R11742	Clone 25-1 encoded

ALIGNMENTS

RESULT 1
R88870
ID R88870 standard; Protein; 359 AA.
XX
AC R88870;
XX
DT 07-NOV-1996 (first entry)
XX
DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW modification; mutation; viral replication; deficient; inhibition;
KW viral resistance; geminivirus; tomato yellow leaf curl virus;
KW Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW AL1 protein; dominant negative phenotype.
XX
OS Sardinian tomato yellow leaf curl virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 227
FT /note= "wild-type Lys has been replaced by Ala"
XX
PN WO9608573-A1.
XX
PD 21-MAR-1996.
XX
PF 15-SEP-1995; 95WO-FR01192.
XX
PR 15-SEP-1994; 94FR-0011040.
XX
PA (CNRS) CENT NAT RECH SCI.
XX
PI Gronenborn B;
XX
DR WPI; 1996-179947/18.

```
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             61.5%;  Score 222;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 5.1e-21;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 kvfqvppap 179


RESULT   2
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             61.5%;  Score 222;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 5.1e-21;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 kvfqvppap 179


RESULT   3
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
```

```
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             61.5%;  Score 222;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 5.1e-21;
  Matches   40;  Conservative   10;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|    |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 kvfqvppap 179


RESULT   4
W34338
ID   W34338 standard; Protein; 353 AA.
XX
AC   W34338;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 protein.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93314.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 219;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.2e-20;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::|||: |:|||| | | || |:    |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   5
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;
```

```
  Query Match             60.7%;  Score 219;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.2e-20;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 219;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.2e-20;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   7
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 219;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.2e-20;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179
```

```
RESULT   8
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 219;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.2e-20;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   9
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             60.1%;  Score 217;  DB 8;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.3e-20;
  Matches   42;  Conservative   11;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |:    |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.8e-20;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::    |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.8e-20;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::    |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.8e-20;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::    |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.8e-20;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::    |: :: | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             56.4%;  Score 203.5;  DB 8;  Length 361;
  Best Local Similarity   63.1%;  Pred. No. 1.4e-18;
  Matches   41;  Conservative    9;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRIF 64
           | |||||||||||  |::||| |||||: || :|| |:    |::|: |||||| || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
           |: |
Db     174 sdevp 178


RESULT  15
W34329
ID   W34329 standard; Protein; 357 AA.
XX
AC   W34329;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato yellow leaf curl virus mutant C1 protein pTYK104R #1.
XX
KW   Geminivirus; pTYK104R #1; C1 protein; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato yellow leaf curl virus strain Israel.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93287.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 4.5; Page 84-86; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the tomato
CC   yellow leaf curl virus (TYLCV) C1 protein (see W34337) that is
CC   required for replication.  It is encoded by construct pTYK104R #1
CC   (see T93287) of an infectious clone of TYLCV and carries a single
CC   mutation in the DNA-nicking domain.  The invention involves
CC   production of transgenic plants containing DNA comprising
CC   geminivirus C1 or AC1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from TYLCV, tomato mottle virus or
CC   bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   357 AA;


  Query Match             55.4%;  Score 200;  DB 18;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 3.9e-18;
  Matches   39;  Conservative    9;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIAAAIPEKYLFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||| |||||: || ||| |:    |: |:|||||:||||||
Db     111 fgvsqidgrsarggqqsandayaealnsgsksealnilkekapkdyilqfhnlssnldri 170

Qy      64 F 64
           |
Db     171 f 171


Search completed: February  3, 2001, 02:15:27
Job time: 3434 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:27 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	220	61.3	361	18	W34336	Tomato mottle viru
2	220	61.3	361	18	W34324	Tomato mottle viru
3	220	61.3	361	18	W34325	Tomato mottle viru
4	220	61.3	361	18	W34326	Tomato mottle viru
5	219	61.0	359	17	R88870	Sardinian tomato y
6	219	61.0	359	17	R88871	Sardinian tomato y
7	219	61.0	359	17	R88872	Sardinian tomato y
8	218	60.7	353	18	W34338	Bean golden mosaic
9	218	60.7	353	18	W34332	Bean golden mosaic
10	218	60.7	353	18	W34333	Bean golden mosaic
11	218	60.7	353	18	W34334	Bean golden mosaic
12	218	60.7	353	18	W34335	Bean golden mosaic
13	216	60.2	353	8	P70407	ORF 4 gene product
14	208.5	58.1	361	8	P70562	Product of ORF 4 f
15	208	57.9	362	19	W56495	Tobacco leaf curli
16	203	56.5	357	18	W34329	Tomato yellow leaf
17	203	56.5	357	18	W34330	Tomato yellow leaf
18	203	56.5	357	18	W34331	Tomato yellow leaf
19	195	54.3	357	18	W34337	Tomato yellow leaf
20	66.5	18.5	512	19	W68473	HIV-1 strain YBF30
21	63.5	17.7	1713	16	R70148	Deduced sequence o
22	60.5	16.9	131	18	W34327	Tomato mottle viru
23	60.5	16.9	447	21	Y87089	Human secreted pro
24	60.5	16.9	447	21	Y66744	Membrane-bound pro
25	60.5	16.9	447	21	Y50947	Human adult aorta
26	60.5	16.9	456	21	Y87190	Human secreted pro
27	59.5	16.6	619	13	R27651	Human calcium chan
28	59.5	16.6	2161	14	R33545	Sequence of the al
29	59.5	16.6	2161	16	R71001	Human neuronal cal
30	59.5	16.6	2161	16	R71002	Human neuronal cal
31	59.5	16.6	2161	19	W63137	Human calcium chan
32	59.5	16.6	2161	19	W63149	Human calcium chan
33	59	16.4	164	20	Y73899	Human prostate tum
34	58.5	16.3	374	19	W37994	Mutant Aspergillus
35	58.5	16.3	374	20	Y39875	A. oryzae P4-8.1 l
36	56.5	15.7	4472	17	R97245	Virulence gene clu
37	56	15.6	62	20	Y11882	Human 5' EST secre
38	56	15.6	292	21	Y92317	Geminivirus Rep C1
39	56	15.6	335	21	Y92318	Mastrevirus Rep in
40	56	15.6	1817	19	W64193	Lettuce resistance
41	55.5	15.5	266	19	W72937	Mycobacterium tube
42	55.5	15.5	266	20	Y21957	Amino acid sequenc
43	55	15.3	362	12	R11468	Prepro-polygalactu
44	55	15.3	433	18	W13406	Solanum melongena
45	54.5	15.2	320	20	Y30730	Amino acid sequenc

ALIGNMENTS

RESULT 1
W34336
ID W34336 standard; Protein; 361 AA.
XX
AC W34336;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 protein.
XX
KW Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93294.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

XX
PS Example 3.3; Page 57-58; 132pp; English.
XX
CC This protein comprises the wild-type AC1 protein of tomato mottle
CC virus (ToMoV), a geminivirus that has a bipartite genome. The AC1
CC gene (see T93294) must be expressed for efficient replication of
CC the two genomic components, DNA-A and DNA-B. The AC1 protein has a
CC DNA binding site specific to the DNA-A common region, a DNA nicking
CC activity, and an NTP binding activity. The invention involves
CC production of transgenic plants containing DNA comprising AC1 or C1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC W34324-35) that have mutations in the highly conserved DNA-nicking
CC domain and/or the NTP-binding domain.
XX
SQ Sequence 361 AA;

Query Match 61.3%; Score 220; DB 18; Length 361;
Best Local Similarity 58.6%; Pred. No. 1.4e-20;
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
|: ||:||:|||||||| |::||: |:|||||| : || ::||: | : | ||: |||
Db 110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy 61 DRIFDKTPEP 70
:||| | |||
Db 170 erifakapep 179

RESULT 2
W34324
ID W34324 standard; Protein; 361 AA.
XX
AC W34324;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93282.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 3.4; Page 60-62; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant, designated
CC ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC W34336). It is encoded by a mutant AC1 gene (see T93282) of
CC tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC NTP-binding domains. The AC1 gene (see also T93294) must be
CC expressed for efficient replication of the two genomic components,
CC DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC involves production of transgenic plants containing DNA comprising
CC geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC interfere in trans with geminiviral replication during infection.
CC Such transgenic plants are resistant to viral infection. The
CC AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC polypeptides (see W34324-35) that have mutations in the highly
CC conserved DNA-nicking and/or NTP-binding domains.
XX
SQ Sequence 361 AA;

Query Match 61.3%; Score 220; DB 18; Length 361;
Best Local Similarity 58.6%; Pred. No. 1.4e-20;
Matches 41; Conservative 14; Mismatches 15; Indels 0; Gaps 0;

Qy 1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
|: ||:||:|||||||| |::||: |:|||||| : || ::||: | : | ||: |||
Db 110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy 61 DRIFDKTPEP 70
:||| | |||
Db 170 erifakapep 179

RESULT 3
W34325
ID W34325 standard; Protein; 361 AA.
XX
AC W34325;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato mottle virus isolate Florida.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93283.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 3.5; Page 64-65; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant, designated
CC ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see

```
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match           61.3%;  Score 220;  DB 18;  Length 361;
  Best Local Similarity   58.6%;  Pred. No. 1.4e-20;
  Matches   41;  Conservative   14;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::||: |   : | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT   4
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match           61.3%;  Score 220;  DB 18;  Length 361;
  Best Local Similarity   58.6%;  Pred. No. 1.4e-20;
  Matches   41;  Conservative   14;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::||: |   : | ||: |||
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifakapep 179


RESULT   5
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
```

```
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             61.0%;  Score 219;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.8e-20;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 lewgtfqidgrsarggqqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   6
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             61.0%;  Score 219;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.8e-20;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 lewgtfqidgrsarggqqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   7
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
```

```
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match            61.0%;  Score 219;  DB 17;  Length 359;
  Best Local Similarity  58.0%;  Pred. No. 1.8e-20;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   8
W34338
ID   W34338 standard; Protein; 353 AA.
XX
AC   W34338;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 protein.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93314.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            60.7%;  Score 218;  DB 18;  Length 353;
  Best Local Similarity  60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   9
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match            60.7%;  Score 218;  DB 18;  Length 353;
  Best Local Similarity  60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
```

```
          :||| | |||
Db    170 erifvkvpep 179


RESULT  10
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 218;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT  11
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 218;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT  12
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
```

```
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 218;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT  13
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             60.2%;  Score 216;  DB 8;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 4.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             58.1%;  Score 208.5;  DB 8;  Length 361;
  Best Local Similarity   66.2%;  Pred. No. 4.3e-19;
  Matches   43;  Conservative    7;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRIF 64
           | ||||||||||| |::||| |||||: || :|| |:||| |   : |||||| || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             57.9%;  Score 208;  DB 19;  Length 362;
  Best Local Similarity   53.6%;  Pred. No. 5e-19;
  Matches   45;  Conservative    8;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | :| |||||||||  ||::||| |||:|: ||  || |:||| |   : ||||||||||
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 RI----------------FDKTPE 69
           ||                ||: ||
Db     171 rifapplevfvcpfssssfdqvpe 194



Search completed: February  3, 2001, 02:15:28
Job time: 3435 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:43 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	219	61.0	359	3	US-08-809-103B-2	Sequence 2, Appli
2	219	61.0	359	3	US-08-809-103B-4	Sequence 4, Appli
3	219	61.0	359	3	US-08-809-103B-6	Sequence 6, Appli
4	219	61.0	359	3	US-08-809-103B-8	Sequence 8, Appli
5	63.5	17.7	1713	3	US-08-600-982-24	Sequence 24, Appl
6	63.5	17.7	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
7	59.5	16.6	2161	1	US-07-745-206A-2	Sequence 2, Appli
8	59.5	16.6	2161	1	US-08-455-543A-49	Sequence 49, Appl
9	59.5	16.6	2161	1	US-08-455-543A-51	Sequence 51, Appl
10	59.5	16.6	2161	2	US-08-223-305C-49	Sequence 49, Appl
11	59.5	16.6	2161	2	US-08-223-305C-51	Sequence 51, Appl
12	59.5	16.6	2161	2	US-08-311-363-2	Sequence 2, Appli
13	58.5	16.3	374	2	US-08-928-692-51	Sequence 51, Appl
14	56	15.6	171	2	US-08-609-049A-22	Sequence 22, Appl
15	54.5	15.2	3898	2	US-08-876-991-2	Sequence 2, Appli
16	54.5	15.2	3898	2	US-09-059-853-2	Sequence 2, Appli
17	54	15.0	242	2	US-08-622-352A-3	Sequence 3, Appli
18	54	15.0	242	3	US-08-826-390-3	Sequence 3, Appli
19	54	15.0	271	1	US-08-276-919-10	Sequence 10, Appl
20	54	15.0	271	1	US-08-776-088-13	Sequence 13, Appl
21	54	15.0	271	4	PCT-US95-09145A-13	Sequence 13, Appl
22	54	15.0	274	1	US-08-776-088-15	Sequence 15, Appl
23	54	15.0	274	4	PCT-US95-09145A-15	Sequence 15, Appl
24	54	15.0	325	1	US-08-276-919-4	Sequence 4, Appli
25	54	15.0	325	1	US-08-276-919-13	Sequence 13, Appl
26	54	15.0	325	1	US-08-776-088-4	Sequence 4, Appli
27	54	15.0	325	1	US-08-776-088-18	Sequence 18, Appl
28	54	15.0	325	4	PCT-US95-09145A-4	Sequence 4, Appli
29	54	15.0	325	4	PCT-US95-09145A-18	Sequence 18, Appl
30	54	15.0	354	1	US-08-276-919-2	Sequence 2, Appli
31	54	15.0	354	1	US-08-776-088-2	Sequence 2, Appli
32	54	15.0	354	1	US-08-776-088-6	Sequence 6, Appli
33	54	15.0	354	3	US-09-325-320-2	Sequence 2, Appli
34	54	15.0	354	4	PCT-US95-09145A-2	Sequence 2, Appli
35	54	15.0	354	4	PCT-US95-09145A-6	Sequence 6, Appli
36	53.5	14.9	127	2	US-08-637-759B-334	Sequence 334, App
37	53.5	14.9	127	3	US-08-871-355A-334	Sequence 334, App
38	53	14.8	266	2	US-07-857-224B-40	Sequence 40, Appl
39	53	14.8	446	2	US-08-672-814D-11	Sequence 11, Appl
40	53	14.8	1298	2	US-08-690-473-2	Sequence 2, Appli
41	52.5	14.6	81	3	US-08-995-159-6	Sequence 6, Appli
42	52.5	14.6	312	1	US-08-414-625-2	Sequence 2, Appli
43	52	14.5	338	2	US-08-674-149A-12	Sequence 12, Appl
44	52	14.5	482	1	US-07-792-885A-1	Sequence 1, Appli
45	52	14.5	482	1	US-08-142-439A-7	Sequence 7, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match 61.0%;  Score 219;  DB 3;  Length 359;
  Best Local Similarity 58.0%;  Pred. No. 2.6e-21;
  Matches 40;  Conservative 11;  Mismatches 18;  Indels 0;  Gaps 0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             61.0%;  Score 219;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.6e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             61.0%;  Score 219;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.6e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             61.0%;  Score 219;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.6e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             17.7%;  Score 63.5;  DB 3;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 12;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQFHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   6
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24
```

```
  Query Match             17.7%;  Score 63.5;  DB 4;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 12;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPAAALFQEHNLNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :|  |      |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRKA----KTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   7
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             16.6%;  Score 59.5;  DB 1;  Length 2161;
  Best Local Similarity   28.2%;  Pred. No. 54;
  Matches   20;  Conservative    5;  Mismatches   17;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKI 45
           ||| :  |:||
Db     767 EEAEEKERKKI 777


RESULT   8
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             16.6%;  Score 59.5;  DB 1;  Length 2161;
  Best Local Similarity   28.2%;  Pred. No. 54;
  Matches   20;  Conservative    5;  Mismatches   17;  Indels   29;  Gaps    1;
```

```
Qy      4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
          |     ||  | ||  :|                              | | ||:|| : |
Db    707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy     35 EEALQIIREKI 45
          ||| :  |:||
Db    767 EEAEEKERKKI 777


RESULT  9
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match            16.6%;  Score 59.5;  DB 1;  Length 2161;
  Best Local Similarity  28.2%;  Pred. No. 54;
  Matches   20;  Conservative    5;  Mismatches   17;  Indels   29;  Gaps    1;

Qy      4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
          |     ||  | ||  :|                              | | ||:|| : |
Db    707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy     35 EEALQIIREKI 45
          ||| :  |:||
Db    767 EEAEEKERKKI 777


RESULT  10
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
```

```
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             16.6%;  Score 59.5;  DB 2;  Length 2161;
  Best Local Similarity   28.2%;  Pred. No. 54;
  Matches   20;  Conservative    5;  Mismatches   17;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKI 45
           ||| :  |:||
Db     767 EEAEEKERKKI 777


RESULT  11
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             16.6%;  Score 59.5;  DB 2;  Length 2161;
  Best Local Similarity   28.2%;  Pred. No. 54;
  Matches   20;  Conservative    5;  Mismatches   17;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKI 45
           ||| :  |:||
Db     767 EEAEEKERKKI 777


RESULT  12
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,363
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-51506
;    TELECOMMUNICATION INFORMATION:
```

```
;       TELEPHONE:  (619)238-0999
;       TELEFAX:  (619)238-0062
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2161 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             16.6%;  Score 59.5;  DB 2;  Length 2161;
  Best Local Similarity   28.2%;  Pred. No. 54;
  Matches   20;  Conservative    5;  Mismatches   17;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKI 45
           ||| :  |:||
Db     767 EEAEEKERKKI 777


RESULT  13
US-08-928-692-51
; Sequence 51, Application US/08928692
; Patent No. 5958727
;  GENERAL INFORMATION:
;    APPLICANT:  Brody, Howard
;    APPLICANT:  Yaver, Deborah S.
;    APPLICANT:  Lamsa, Michael
;    APPLICANT:  Hansen, Kim
;    TITLE OF INVENTION:  Methods for Modifying the Production of
;    TITLE OF INVENTION:  a Polypeptide
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,692
;      FILING DATE:  12-SEPT-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4944.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  374 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5958727e
US-08-928-692-51


  Query Match             16.3%;  Score 58.5;  DB 2;  Length 374;
  Best Local Similarity   45.5%;  Pred. No. 7.2;
  Matches   15;  Conservative    5;  Mismatches   10;  Indels    3;  Gaps    1;

Qy      22 NDAAAEALNASSKE---EALQIIREKIPAAALF 51
           :||  ||| ||:|    | ||:| | : ||  :
Db     181 SDAQREALTASAKAAGLEVLQLIHEPVAAALAY 213


RESULT  14
US-08-609-049A-22
; Sequence 22, Application US/08609049A
; Patent No. 5948664
;  GENERAL INFORMATION:
;    APPLICANT:  Williams, Lewis T.
;    APPLICANT:  Molz, Lisa
;    APPLICANT:  Chen, Yen-Wen
;    TITLE OF INVENTION:  No. 5948664el PI 3-Kinase Polypeptides
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:, Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/609,049A
;      FILING DATE:  29-FEB-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dow, Karen B.
;      REGISTRATION NUMBER:  29,684
;      REFERENCE/DOCKET NUMBER:  2307K-063700US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  171 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-609-049A-22


  Query Match             15.6%;  Score 56;  DB 2;  Length 171;
  Best Local Similarity   23.4%;  Pred. No. 5.5;
  Matches   18;  Conservative   14;  Mismatches   33;  Indels   12;  Gaps    2;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAA--EALNASSKEEALQIIREKIPAAAL--------- 50
           |:| |  :| |    ||  : | :   | :: |     :|:    : |||
Db      24 LIWKEAGLDIRMLPYGCLATGDRSGLIEVVSTSETIADIQLNSSNVAAAAAAFNKDALLN 83

Qy      51 -FQFHNLNSNLDRIFDK 66
            : :|   :|||  ::
Db      84 WLKEYNSGDDLDRAIEE 100


RESULT  15
US-08-876-991-2
; Sequence 2, Application US/08876991
; Patent No. 5925360
;  GENERAL INFORMATION:
;    APPLICANT:  Gregor Meyers, Tillmann R menapf,
;    APPLICANT:  Heinz-J rgen Thiel
;    TITLE OF INVENTION:  Hog cholera virus vaccine and diagnostic
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Organon Teknika Corporation
;      ADDRESSEE:  Biotechnology Research Institute
```

```
;       STREET:  1330-A Piccard Drive
;       CITY:  Rockville
;       STATE:  Maryland
;       COUNTRY:  U.S.A.
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/876,991
;       FILING DATE:  16-JUN-1997
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/747,577
;       FILING DATE:
;       APPLICATION NUMBER:  US/08/650,584
;       FILING DATE:
;       APPLICATION NUMBER:  US/08/469,702
;       FILING DATE:
;       APPLICATION NUMBER:  US/08/123,596
;       FILING DATE:
;       APPLICATION NUMBER:  07/797,554
;       FILING DATE:  22-NOV-1991
;       APPLICATION NUMBER:  US 07/494,991
;       FILING DATE:  16-MAR-1990
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  William M. Blackstone
;       REGISTRATION NUMBER:  29,772
;       REFERENCE/DOCKET NUMBER:
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (301) 258-5200
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3898 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-876-991-2


  Query Match             15.2%;  Score 54.5;  DB 2;  Length 3898;
  Best Local Similarity   21.4%;  Pred. No. 5.5e+02;
  Matches   18;  Conservative   15;  Mismatches   34;  Indels   17;  Gaps    4;

Qy       1 TLVWGEF----QVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQ--FH 54
           |  | |:    |:|  : :  | |:      |||  |:     : :  :| :  |:  |
Db     716 TTTWKEYSHGLQLDDGTVKAVC-TAGSFKVTALNVVSRRYLASLHKRALPTSVTFELLFD 774

Qy      55 NLNSNLDRI----------FDKTP 68
             |  :: :          || :|
Db     775 GTNPAIEEMDDDFGFGLCPFDTSP 798


Search completed: February  3, 2001, 02:19:45
Job time: 1707 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:26 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	343	95.5	352	1	VAL1_TGMV	P03567	tomato gold
2	247	68.8	361	1	VAL1_PYMVV	P27258	potato yell
3	243	67.7	358	1	VAL1_CLVK	P14982	cassava lat
4	243	67.7	358	1	VAL1_CLVN	P14972	cassava lat
5	232	64.6	362	1	VAL1_TYLCA	P36279	tomato yell
6	226	63.0	359	1	VAL1_TYLCU	P38609	tomato yell
7	224	62.4	349	1	VAL1_PHUV	Q06923	pepper huas
8	219	61.0	359	1	VAL1_TYLCM	P27260	tomato yell
9	216	60.2	353	1	VAL1_BGMV	P05175	bean golden
10	216	60.2	361	1	VAL1_TMOV	Q06657	tomato mott
11	215	59.9	355	1	VAL1_ABMVW	P21947	abutilon mo
12	212	59.1	358	1	VAL1_BCTV	P14991	beet curly
13	203	56.5	357	1	VAL1_TYLCV	P27259	tomato yell
14	118	32.9	347	1	VAL1_SLCV	P29048	squash leaf
15	65	18.1	630	1	VE1_HPV66	Q80957	human papil
16	63.5	17.7	1713	1	LMA3_HUMAN	Q16787	homo sapien
17	62.5	17.4	428	1	GBA1_CANAL	P28868	candida alb
18	61	17.0	631	1	VE1_HPV30	Q05112	human papil
19	60.5	16.9	1070	1	P11B_HUMAN	P42338	homo sapien
20	59.5	16.6	706	1	Y006_RICPR	Q9zed6	rickettsia
21	59.5	16.6	1053	1	UBA3_WHEAT	P31252	triticum ae
22	59.5	16.6	1610	1	CCAD_MESAU	Q99244	mesocricetu
23	59.5	16.6	2161	1	CCAD_HUMAN	Q01668	homo sapien
24	59.5	16.6	2203	1	CCAD_RAT	P27732	rattus norv
25	59	16.4	703	1	UGS2_RAT	P17625	rattus norv
26	59	16.4	1852	1	CCAS_CYPCA	P22316	cyprinus ca
27	58	16.2	387	1	Y4PF_RHISN	P55615	rhizobium s
28	58	16.2	1608	1	HLYA_SERMA	P15320	serratia ma
29	57.5	16.0	367	1	LHX4_MOUSE	P53776	mus musculu
30	57	15.9	355	1	CRTB_RHOSH	P54905	rhodobacter
31	57	15.9	524	1	MPPA_RAT	P20069	rattus norv
32	57	15.9	874	1	SLAP_BACLI	P49052	bacillus li
33	56	15.6	129	1	RK12_PORPU	P51339	porphyra pu
34	56	15.6	329	1	TC3A_CAEEL	P34257	caenorhabdi
35	55.5	15.5	266	1	ETFB_MYCLE	O33095	mycobacteri
36	55.5	15.5	266	1	ETFB_MYCTU	O53276	mycobacteri
37	55.5	15.5	2190	1	CCAD_CHICK	O73700	gallus gall
38	55	15.3	111	1	RL22_ACHLA	P29222	acholeplasm
39	55	15.3	268	1	IL1B_RABIT	P14628	oryctolagus
40	55	15.3	362	1	PGL2_ASPTU	P19805	aspergillus
41	55	15.3	408	1	CINA_THEMA	Q9s5x1	thermotoga
42	55	15.3	433	1	UFOG_SOLME	Q43641	solanum mel
43	55	15.3	480	1	SYE_HAEIN	P43818	haemophilus
44	55	15.3	525	1	MPPA_HUMAN	Q10713	homo sapien
45	55	15.3	576	1	HEMA_SENDH	P03425	sendai viru

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.5%; Score 343; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 9.5e-32;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991)..
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             68.8%;  Score 247;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 7.8e-21;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| || ||||.|||||: :|| |::||:||:|   |||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             67.7%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 2.2e-20;
  Matches   45;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :| | : |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             67.7%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 2.2e-20;
  Matches   45;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :| | : |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             64.6%;  Score 232;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 4e-19;
  Matches   47;  Conservative    8;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || |||||:||||||||| |::||| |:||| || ||| ::|| |   : |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             63.0%;  Score 226;  DB 1;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 1.9e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             62.4%;  Score 224;  DB 1;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 3.1e-18;
  Matches   43;  Conservative   11;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: |||||:|||||||| |::||  |:|||::| |||||||:|: |     ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             61.0%;  Score 219;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.2e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV     STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.2%;  Score 216;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.5e-17;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: ||:|||||||||||| |::||: |:|||| | | || |::|: |   : | ||: |||
Db    110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFFKVPEP 179


RESULT  10
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             60.2%;  Score 216;  DB 1;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 2.6e-17;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |   : | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             59.9%;  Score 215;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 3.3e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |  |||||:|||||||| ||:||: |:|||   : || |::|: |   : | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  12
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.1%;  Score 212;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 7.3e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |     | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             56.5%;  Score 203;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 7.7e-16;
  Matches   40;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           :|  |:|||||||| |::||| ||||| || ||| |::|| |   : |||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             32.9%;  Score 118;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 3.3e-06;
  Matches   25;  Conservative   12;  Mismatches   25;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |   : :|||| :|:::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
VE1_HPV66
ID   VE1_HPV66      STANDARD;      PRT;   630 AA.
AC   Q80957;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   REPLICATION PROTEIN E1.
GN   E1.
OS   Human papillomavirus type 66.
OC   Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Delius H.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC       VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC       PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC       CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U31794; AAA79501.1; -.
DR   INTERPRO; IPR001177; -.
DR   PFAM; PF00519; E1; 1.
DR   PFAM; PF00524; E1_N; 1.
KW   Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW   Nuclear protein.
FT   NP_BIND     458    465       ATP (POTENTIAL).
SQ   SEQUENCE   630 AA;  72065 MW;  22DDDA5934F7291B CRC64;


  Query Match             18.1%;  Score 65;  DB 1;  Length 630;
  Best Local Similarity   25.5%;  Pred. No. 6.8;
  Matches   13;  Conservative   14;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      19 QTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRIFDKTPE 69
             :|   : :  ||   ::  :|  : : || ||: :    ||:  :: :|||
Db     269 KTITKSLSSILNVPQEQMLIQPPKLRSPAVALYFYKTAMSNISEVYGETPE 319
```

Search completed: February 3, 2001, 02:24:27
Job time: 639 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:38 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	343	95.5	352	1	QQCVL1	AL1 protein - toma
2	247	68.8	361	1	QQCVPT	AL1 protein - pota
3	243	67.7	358	2	S07594	hypothetical prote
4	232	64.6	362	1	JQ1887	AL1 protein - toma
5	226	63.0	359	2	S39211	gene C1 protein -
6	224	62.4	349	2	JQ2300	replicase - pepper
7	224	62.4	349	2	S31875	AL1 protein - pepp
8	219	61.0	359	2	S22593	hypothetical prote
9	216	60.2	351	2	JQ2327	AL1 protein - Indi
10	216	60.2	358	1	JQ1870	AL1 protein - toma
11	215	59.9	355	1	QQCVW1	AV1 protein - abut
12	213	59.3	359	2	S39235	gene C1 protein -
13	212	59.1	385	2	S28360	AL1 protein - beet
14	209	58.2	360	2	S59885	replication-associ
15	203	56.5	357	1	QQCVC1	AL1 protein - toma
16	134	37.3	131	2	S45059	AC1 protein (clone
17	118	32.9	347	1	QQCVS1	AL1 protein - squa
18	65	18.1	587	2	JC1419	Fc gamma (IgG) rec
19	63.5	17.7	1713	2	A55347	adhesive ligand ep
20	62	17.3	160	2	G82060	hypothetical prote
21	61	17.0	631	2	S36505	E1 protein - human
22	61	17.0	840	2	T36175	probable large ATP
23	60.5	16.9	429	2	A44384	GTP-binding regula
24	60.5	16.9	435	2	E59096	hypothetical prote
25	60.5	16.9	447	2	T12544	hypothetical prote
26	60.5	16.9	1070	1	A54600	1-phosphatidylinos
27	60	16.7	397	2	B71078	probable NADH oxid
28	59.5	16.6	706	2	H71707	hypothetical prote
29	59.5	16.6	1053	2	T06483	probable ubiquitin
30	59.5	16.6	1610	2	A46227	voltage-dependent
31	59.5	16.6	1646	2	JH0422	voltage-dependent
32	59.5	16.6	2137	2	T05244	hypothetical prote
33	59.5	16.6	2161	2	JH0564	calcium channel al
34	59.5	16.6	2181	2	A38198	calcium channel al
35	59.5	16.6	2203	2	T42742	voltage-dependent
36	59	16.4	451	2	S26839	retrovirus-related
37	59	16.4	703	2	A35362	UDPglucose--glycog
38	59	16.4	1852	2	A37860	calcium channel pr
39	58.5	16.3	388	2	C69196	hypothetical prote
40	58	16.2	1127	2	G71274	hypothetical prote
41	58	16.2	1245	2	E83110	exodeoxyribonuclea
42	58	16.2	1341	2	JG0166	LaMDR1 protein - L
43	58	16.2	1608	2	A28182	hemolysin A - Serr
44	57.5	16.0	230	2	S46332	homeotic protein G
45	57.5	16.0	392	2	T45290	acetyl-CoA C-acety

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.5%; Score 343; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.1e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||   ||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             68.8%;  Score 247;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 1.1e-20;
  Matches   47;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: || ||:|||||||| || |||||||||: :|| |::||:||:|   |||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             67.7%;  Score 243;  DB 2;  Length 358;
  Best Local Similarity   64.3%;  Pred. No. 3.1e-20;
  Matches   45;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: ||:||:|||||||| |::||| |:||: || ||| :||| :|   : |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             64.6%;  Score 232;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 5.6e-19;
  Matches   47;  Conservative    8;  Mismatches   14;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          || |||||:|||||||| |::||| |:|||  || ||| ::||  |   : |||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.0%;  Score 226;  DB 2;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 2.7e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.4%;  Score 224;  DB 2;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 4.5e-18;
  Matches   43;  Conservative   11;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: |||||:|||||||| |::||  |:|||::| |||||||:|: | |    ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.4%;  Score 224;  DB 2;  Length 349;
  Best Local Similarity   61.4%;  Pred. No. 4.5e-18;
  Matches   43;  Conservative   11;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: |||||:|||||||| |::|| |:|||::| ||||||:|: |     ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.0%;  Score 219;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.7e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |   :  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 2;  Length 351;
  Best Local Similarity   62.7%;  Pred. No. 3.7e-17;
  Matches   42;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
           || ||:|||||||| |::||| | |||: || |||:|:||  |   |  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 1;  Length 358;
  Best Local Similarity   57.1%;  Pred. No. 3.8e-17;
  Matches   40;  Conservative   14;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||:  |   : | ||: |||
Db     107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     167 ERIFAKAPEP 176


RESULT  11
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.9%;  Score 215;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 4.9e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;
```

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |  |||||:||||||||| ||:||: |:||||    : || }::|: |    : | ||: |||
Db    110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179


RESULT  12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.3%;  Score 213;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 8.4e-17;
  Matches   39;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|::  |    :  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.1%;  Score 212;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.2e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |     | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
           :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 209;  DB 2;  Length 360;
  Best Local Similarity   62.1%;  Pred. No. 2.4e-16;
  Matches   41;  Conservative   10;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
          :| ||:|||||||| |::||| |||:|: ||  || |:||| |   : ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178


RESULT  15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.5%;  Score 203;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 1.2e-15;
  Matches   40;  Conservative    9;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
          :|  |:|||||||| |::||| |||||: || ||| |::|| |   : ||||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


Search completed: February  3, 2001, 02:17:39
Job time: 2211 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:51 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-5
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	287	79.9	226	12	Q9WHF6	Q9whf6 tomato mild
2	277	77.2	361	12	Q67574	bean golden
3	275	76.6	225	12	O9QDB1	Q9qdb1 cowpea gold
4	275	76.6	226	12	O09727	O09727 leonurus mo
5	271	75.5	185	12	Q98693	Q98693 sida golden
6	271	75.5	364	12	Q9QS55	Q9qs55 sweet potat
7	264	73.5	149	12	P88975	P88975 macroptiliu
8	264	73.5	190	12	Q9Z089	Q9z089 tobacco lea
9	264	73.5	190	12	Q9Z084	Q9z084 tobacco lea
10	264	73.5	233	12	Q9YLA4	Q9yla4 macroptiliu
11	263	73.3	190	12	Q9W827	Q9w827 tobacco lea
12	263	73.3	208	12	Q9Z0C4	Q9z0c4 tobacco lea
13	261	72.7	208	12	Q9Z0C0	Q9z0c0 tobacco lea
14	261	72.7	208	12	Q9Z0B8	Q9z0b8 tobacco lea
15	259	72.1	208	12	Q9Z0B6	Q9z0b6 tobacco lea
16	258	71.9	203	12	Q9Z083	Q9z083 tobacco lea
17	258	71.9	234	12	O39180	O39180 potato yell
18	251	69.9	208	12	Q9Z0A0	Q9z0a0 tobacco lea
19	251	69.9	363	12	O72705	O72705 cotton leaf
20	251	69.9	363	12	O72719	O72719 cotton leaf
21	250	69.6	190	12	Q9Z086	Q9z086 tobacco lea
22	250	69.6	208	12	Q9Z0C6	Q9z0c6 tobacco lea
23	249	69.4	349	12	Q88888	Q88888 tomato pseu
24	249	69.4	363	12	O73577	O73577 cotton leaf
25	246	68.5	190	12	Q9Z0A7	Q9z0a7 tobacco lea
26	246	68.5	362	12	Q9IN48	Q9in48 okra enatio
27	245	68.2	307	12	Q9IET7	Q9iet7 cotton leaf
28	245	68.2	361	12	O72723	O72723 cotton leaf
29	245	68.2	362	12	O56816	O56816 chayote mos
30	244	68.0	231	12	Q96620	Q96620 african tom
31	244	68.0	354	12	Q9IN42	Q9in42 south afric
32	244	68.0	359	12	Q9YZV4	Q9yzv4 tomato yell
33	244	68.0	359	12	Q9YZV2	Q9yzv2 tomato yell
34	244	68.0	359	12	Q9YUX7	Q9yux7 tomato yell
35	244	68.0	359	12	Q9YL27	Q9yl27 tomato yell
36	243	67.7	358	12	Q9JEA2	Q9jea2 cassava gem
37	243	67.7	359	12	Q88942	Q88942 tomato yell
38	239	66.6	362	12	Q9YYT7	Q9yyt7 althea rose
39	239	66.6	363	12	O73494	O73494 okra yellow
40	237	66.0	358	12	Q9WR17	Q9wr17 african cas
41	237	66.0	358	12	Q9JEA9	Q9jea9 cassava gem
42	236	65.7	360	12	O72707	O72707 cotton leaf
43	236	65.7	360	12	O72710	O72710 cotton leaf
44	235	65.5	153	12	O41475	O41475 wissadula g
45	235	65.5	361	12	Q9QPQ0	Q9qpq0 tomato leaf

ALIGNMENTS

RESULT 1
Q9WHF6
ID Q9WHF6 PRELIMINARY; PRT; 226 AA.
AC Q9WHF6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN REP.
OS tomato mild mottle geminivirus.
OC Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX NCBI_TaxID=92943;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HN96-H5KW;
RA Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA Maxwell D.P.;
RT "Molecular characterization and DNA-based detection methods for
RT vegetable-infecting geminiviruses in Central America.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF131071; AAD33471.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25941 MW; 2EA4116712871A23 CRC64;

Query Match 79.9%; Score 287; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 7.3e-26;
Matches 54; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: ||||:|||||||| ||:|||||||||||||||::||:||:|   |||:|||:|||
Db    111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180
```

```
RESULT   2
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.2%;  Score 277;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 1.8e-24;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
          || |||||||||||  ||:||||:|||||||||||:|||:||:|   |||:|||:||||||
Db    113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |:|
Db    173 FTKAPDP 179


RESULT   3
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.6%;  Score 275;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 1.8e-24;
  Matches   52;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNLDRI 63
          || |||||||||  || ||||:|||||||||||:|||:||:|   |||:|||:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT   4
O09727
ID   O09727      PRELIMINARY;      PRT;   226 AA.
AC   O09727;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   REP.
OS   Leonurus mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=58177;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LEMV- BRAZIL 1;
RA   Faria J.C., Maxwell D.P.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U92532; AAB51157.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25617 MW;  73CDB6E766083FC5 CRC64;


  Query Match             76.6%;  Score 275;  DB 12;  Length 226;
  Best Local Similarity   77.1%;  Pred. No. 1.8e-24;
  Matches   54;  Conservative    4;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
          |: ||||||||||:||| || ||||||||||  |  |||||:||:|  ||||||||||
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          |||| |:|||
Db    171 DRIFAKAPEP 180


RESULT   5
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.5%;  Score 271;  DB 12;  Length 185;
  Best Local Similarity   71.4%;  Pred. No. 4.3e-24;
  Matches   50;  Conservative   11;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:|||||:|   |||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   6
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.5%;  Score 271;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 9.3e-24;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: ||||||||||||||| ||:|||||||||| ||| ||||||||:|   ||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 149;
  Best Local Similarity   70.0%;  Pred. No. 2.2e-23;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|   |||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT   8
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 2.9e-23;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || ||:|||||||||| :||| ||||||||| ||| |||||:|   :||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144
```

```
Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT   9
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 2.9e-23;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
             || || ||:|||||||||| :||| |||||||| ||| ||||:|   :||:|||||||
Db      85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  10
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.5%;  Score 264;  DB 12;  Length 233;
  Best Local Similarity   70.0%;  Pred. No. 3.7e-23;
  Matches   49;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
             |: || ||:|||||||| ||||||||||||: :|| |::|::||:|   |||:|||:|||
Db     110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     170 DRIFMKDPEP 179


RESULT  11
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             73.3%;  Score 263;  DB 12;  Length 190;
  Best Local Similarity   62.4%;  Pred. No. 3.8e-23;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
             || || |||||||||||||| :||| ||||||:|| ||| |||||:|   :||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  12
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
```

```
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             73.3%;  Score 263;  DB 12;  Length 208;
  Best Local Similarity   62.4%;  Pred. No. 4.2e-23;
  Matches   53;  Conservative    6;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || |||||||||||||| :||| |||||||:|| ||| |||||:|   :||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  13
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             72.7%;  Score 261;  DB 12;  Length 208;
  Best Local Similarity   61.2%;  Pred. No. 7.3e-23;
  Matches   52;  Conservative    7;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||| :|| |||||:|   :||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  14
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;


  Query Match             72.7%;  Score 261;  DB 12;  Length 208;
  Best Local Similarity   61.2%;  Pred. No. 7.3e-23;
  Matches   52;  Conservative    7;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||| :|| |||||:|   :||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B6
ID   Q9Z0B6      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B6;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-3(A055),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGA055;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001298; BAA34000.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
```

```
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23586 MW;  A5E61180BF05B9BE CRC64;


  Query Match             72.1%;  Score 259;  DB 12;  Length 208;
  Best Local Similarity   60.0%;  Pred. No. 1.2e-22;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPAAALFQFHNLNSNL 60
           || || ||:|||||||||| :||| |||||||::| ||| |||||:|   :||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNAATKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFSPPLEVFVCPFSSSSFDQVPE 186


Search completed: February  3, 2001, 02:22:51
Job time: 1888 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:51 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	272	76.2	226	12	O09727	O09727 leonurus mo
2	272	76.2	226	12	Q9WHF6	Q9whf6 tomato mild
3	270	75.6	361	12	Q67574	Q67574 bean golden
4	268	75.1	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	268	75.1	364	12	Q9QS55	Q9qs55 sweet potat
6	265	74.2	185	12	Q98693	Q98693 sida golden
7	257	72.0	149	12	P88975	P88975 macroptiliu
8	257	72.0	233	12	Q9YLA4	Q9yla4 macroptiliu
9	251	70.3	234	12	O39180	O39180 potato yell
10	248	69.5	190	12	Q9Z089	Q9z089 tobacco lea
11	248	69.5	190	12	Q9Z084	Q9z084 tobacco lea
12	247	69.2	190	12	Q9W827	Q9w827 tobacco lea
13	247	69.2	208	12	Q9Z0C4	Q9z0c4 tobacco lea
14	245	68.6	208	12	Q9Z0C0	Q9z0c0 tobacco lea
15	245	68.6	208	12	Q9Z0B8	Q9z0b8 tobacco lea
16	243	68.1	208	12	Q9Z0B6	Q9z0b6 tobacco lea
17	242	67.8	203	12	Q9Z083	Q9z083 tobacco lea
18	242	67.8	349	12	Q88888	Q88888 tomato pseu
19	242	67.8	363	12	O73577	O73577 cotton leaf
20	239	66.9	208	12	Q9Z0A0	Q9z0a0 tobacco lea
21	238	66.7	208	12	Q9Z0C6	Q9z0c6 tobacco lea
22	235	65.8	363	12	O72705	O72705 cotton leaf
23	235	65.8	363	12	O72719	O72719 cotton leaf
24	234	65.5	190	12	Q9Z086	Q9z086 tobacco lea
25	233	65.3	362	12	O56816	O56816 chayote mos
26	232	65.0	359	12	Q9YZV4	Q9yzv4 tomato yell
27	232	65.0	359	12	Q9YZV2	Q9yzv2 tomato yell
28	232	65.0	359	12	Q9YUX7	Q9yux7 tomato yell
29	232	65.0	359	12	Q9YL27	Q9yl27 tomato yell
30	231	64.7	353	12	O72692	O72692 beet curly
31	231	64.7	354	12	Q65438	Q65438 beet curly
32	231	64.7	359	12	Q88942	Q88942 tomato yell
33	230	64.4	190	12	Q9Z0A7	Q9z0a7 tobacco lea
34	230	64.4	362	12	Q9IN48	Q9in48 okra enatio
35	229	64.1	307	12	Q9IET7	Q9iet7 cotton leaf
36	229	64.1	361	12	O72723	O72723 cotton leaf
37	228	63.9	231	12	Q96620	Q96620 african tom
38	228	63.9	354	12	Q9IN42	Q9in42 south afric
39	227	63.6	358	12	Q65418	Q65418 beet curly
40	227	63.6	358	12	Q9JEA2	Q9jea2 cassava gem
41	223	62.5	362	12	Q9YYT7	Q9yyt7 althea rose
42	223	62.5	363	12	O73494	O73494 okra yellow
43	221	61.9	306	12	Q9IET1	Q9iet1 cotton leaf
44	221	61.9	358	12	Q9WR17	Q9wr17 african cas
45	221	61.9	358	12	Q9JEA9	Q9jea9 cassava gem

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.2%; Score 272; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 6.7e-23;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||||||:||| || |||||||||| | |||||:||:|||||| | |||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             76.2%;  Score 272;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 6.7e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: |||||:|||||||| ||:|||||||||||||||||::||:||:|:::|| :  |:|||
Db    111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             75.6%;  Score 270;  DB 12;  Length 361;
  Best Local Similarity   76.1%;  Pred. No. 1.9e-22;
  Matches   51;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          || |||||||||| ||:||||:||||||||||:|||:||:|||:|| :  |:||||||
Db    113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |:|
Db    173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             75.1%;  Score 268;  DB 12;  Length 225;
  Best Local Similarity   76.1%;  Pred. No. 1.9e-22;
  Matches   51;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          || ||:|||||||| || ||||:||||||||||:|||:||:|||:|| :  |:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.1%;  Score 268;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 3.3e-22;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||||||||||| ||:|||||||||| ||| ||||||||:|||||| |  | |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             74.2%;  Score 265;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 3.2e-22;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| ||:|||||||||: :||:||:|||||:||:||| :  [:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             72.0%;  Score 257;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 2e-21;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|| :  |:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             72.0%;  Score 257;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 3.3e-21;
```

```
  Matches   48;  Conservative   12;  Mismatches   10;  Indels   0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: || ||:|||||||| |||||||||||: :|| |::|::||:|||:|| :  |:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   9
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             70.3%;  Score 251;  DB 12;  Length 234;
  Best Local Similarity   67.1%;  Pred. No. 1.6e-20;
  Matches   47;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: || ||:|||||||| ||:|||||||||: :||||::||:||:|||:|| :  |:|:|
Db    110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy     61 DRIFDKTPEP 70
          |||| | | |
Db    170 DRIFMKAPNP 179


RESULT  10
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             69.5%;  Score 248;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.7e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          || || ||:|||||||||| :||| ||||||||| ||| |||||:|:|:::| :  |||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             69.5%;  Score 248;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 2.7e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          || || ||:|||||||||| :||| ||||||||| ||| |||||:|:|:::| :  |||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             69.2%;  Score 247;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 3.5e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || |||||||||||||| :||| ||||||:|| ||| |||||:|: ::| :  |||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  13
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             69.2%;  Score 247;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 3.9e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || |||||||||||||| :||| ||||||:|| ||| |||||:|: ::| :  |||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  14
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             68.6%;  Score 245;  DB 12;  Length 208;
  Best Local Similarity   57.6%;  Pred. No. 6.5e-20;
  Matches   49;  Conservative    9;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || ||:|||||||||| :||| ||||||||| :|| |||||:|: ::| :  |||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;
```

Query Match 68.6%; Score 245; DB 12; Length 208;
Best Local Similarity 57.6%; Pred. No. 6.5e-20;
Matches 49; Conservative 9; Mismatches 11; Indels 16; Gaps 1;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || || ||:|||||||||| :||| |||||||| :|| |||||:|: ::| :  |||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186
```

Search completed: February 3, 2001, 02:22:51
Job time: 1888 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:27 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	340	95.2	352	1	VAL1_TGMV	P03567 tomato gold
2	240	67.2	361	1	VAL1_PYMVV	P27258 potato yell
3	227	63.6	358	1	VAL1_CLVK	P14982 cassava lat
4	227	63.6	358	1	VAL1_CLVN	P14972 cassava lat
5	220	61.6	362	1	VAL1_TYLCA	P36279 tomato yell
6	217	60.8	359	1	VAL1_TYLCU	P38609 tomato yell
7	210	58.8	359	1	VAL1_TYLCM	P27260 tomato yell
8	209	58.5	349	1	VAL1_PHUV	Q06923 pepper huas
9	204	57.1	353	1	VAL1_BGMV	P05175 bean golden
10	203	56.9	355	1	VAL1_ABMVW	P21947 abutilon mo
11	200	56.0	358	1	VAL1_BCTV	P14991 beet curly
12	200	56.0	361	1	VAL1_TMOV	Q06657 tomato mott
13	191	53.5	357	1	VAL1_TYLCV	P27259 tomato yell
14	104	29.1	347	1	VAL1_SLCV	P29048 squash leaf
15	64.5	18.1	1713	1	LMA3_HUMAN	Q16787 homo sapien
16	63.5	17.8	1610	1	CCAD_MESAU	Q99244 mesocricetu
17	63.5	17.8	2161	1	CCAD_HUMAN	Q01668 homo sapien
18	63.5	17.8	2203	1	CCAD_RAT	P27732 rattus norv
19	60	16.8	269	1	ORN_YEAST	P54964 saccharomyc
20	60	16.8	295	1	VAL1_TYDVA	P31617 tobacco yel
21	60	16.8	355	1	CRTB_RHOSH	P54905 rhodobacter
22	59.5	16.7	299	1	Y175_HELPJ	Q9zmq7 helicobacte
23	59.5	16.7	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
24	59.5	16.7	4385	1	YP73_CAEEL	Q09222 caenorhabdi
25	58.5	16.4	706	1	Y006_RICPR	Q9zed6 rickettsia
26	58.5	16.4	1232	1	YOQ5_CAEEL	P34643 caenorhabdi
27	58	16.2	207	1	IL6_MARMO	O35736 marmota mon
28	58	16.2	447	1	CDSA_DROME	P56079 d phosphati
29	57.5	16.1	367	1	LHX4_MOUSE	P53776 mus musculu
30	57.5	16.1	844	1	SECA_STACA	P47994 staphylococ
31	57.5	16.1	1852	1	CCAS_CYPCA	P22316 cyprinus ca
32	57	16.0	1483	1	CYP1_YEAST	P12351 saccharomyc
33	57	16.0	1608	1	HLYA_SERMA	P15320 serratia ma
34	56.5	15.8	2190	1	CCAD_CHICK	O73700 gallus gall
35	56	15.7	258	1	YL48_MYCTU	O06228 mycobacteri
36	56	15.7	408	1	CINA_THEMA	Q9s5x1 thermotoga
37	56	15.7	703	1	UGS2_RAT	P17625 rattus norv
38	56	15.7	767	1	TOP1_MOUSE	Q04750 mus musculu
39	56	15.7	1033	1	YDK9_SCHPO	P87115 schizosacch
40	56	15.7	1051	1	CARB_SULSO	Q59969 sulfolobus
41	56	15.7	1663	1	HAPP_HUMAN	O60229 homo sapien
42	55.5	15.5	266	1	ETFB_MYCLE	O33095 mycobacteri
43	55.5	15.5	266	1	ETFB_MYCTU	O53276 mycobacteri
44	55.5	15.5	419	1	Y223_MYCPN	P75465 mycoplasma
45	55.5	15.5	967	1	HEPA_ECOLI	P23852 escherichia

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.2%; Score 340; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 7.8e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             67.2%;  Score 240;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 1.2e-19;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: || ||:||||||||| || |||||||||: :|| |::||:||:|||:|| :  |: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             63.6%;  Score 227;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 3.4e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: ::  |  |||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             63.6%;  Score 227;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 3.4e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: ::  |  |||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             61.6%;  Score 220;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 2.1e-17;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           || |||||:||||||||  |::||| |:|||  || ||| ::||  |: |:  |  |||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             60.8%;  Score 217;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 4.5e-17;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
```

```
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             58.8%;  Score 210;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.8e-16;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   8
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             58.5%;  Score 209;  DB 1;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 3.5e-16;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||||:|||||||| |::||  |:|||::| |||||||:|: |: :   |  : ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             57.1%;  Score 204;  DB 1;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-15;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:|||||||||| |::||: |:|||| | | || |::|: |: ||:     : |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             56.9%;  Score 203;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 1.7e-15;
  Matches   39;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |  |||||:||||||| ||:||: |:|||| : || |::|: |: |:     : |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             56.0%;  Score 200;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 3.6e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||| ||:||: |:|||| : || |::|: |: |       | :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             56.0%;  Score 200;  DB 1;  Length 361;
  Best Local Similarity   52.9%;  Pred. No. 3.6e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: ::     : |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             53.5%;  Score 191;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 3.7e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          :|  |:||||||| |::||| |||||: || ||| |::|| |: |:  |  |:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             29.1%;  Score 104;  DB 1;  Length 347;
  Best Local Similarity   34.8%;  Pred. No. 0.00019;
  Matches   23;  Conservative   14;  Mismatches   25;  Indels    4;  Gaps    1;

Qy      5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRIF 64
          |:::| |    |     :|    |:|| |  ||| ||:   |: :: | :  | :|::|:|
Db    116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy     65 DKTPEP 70
          | |||
Db    172 QKPPEP 177


RESULT  15
LMA3_HUMAN
ID   LMA3_HUMAN     STANDARD;      PRT;  1713 AA.
AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   LAMININ ALPHA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KERATINOCYTES;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC       VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC       6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC       PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC       KERATINOCYTES.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC       NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC       AND LAMININ-7 (KS-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC       TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
```

```
CC       A.
CC   -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC       EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC       SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC       AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC       DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC       IN THE CNS.
CC   -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN G IS GLOBULAR.
CC   -!- DISEASE: JUNCTIONAL EPIDERMOLYSIS BULLOSA GRAVIS IS A BLISTERING
CC       DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION OF BASAL
CC       CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED NUMBER OF
CC       HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT MEMBRANE OF
CC       PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS BULLOSA.
CC   -!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   HSSP; P02468; 1TLE.
DR   MIM; 600805; -.
DR   MIM; 226700; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR001791; -.
DR   INTERPRO; IPR002049; -.
DR   PFAM; PF00053; laminin_EGF; 2.
DR   PFAM; PF00054; laminin_G; 2.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     20       POTENTIAL.
FT   CHAIN       21   1713       LAMININ ALPHA-3 CHAIN.
FT   DOMAIN      46    201       DOMAIN III A.
FT   DOMAIN      67    185       2.5 X LAMININ EGF-LIKE REPEATS.
FT   DOMAIN      67    113       LAMININ EGF-LIKE 1.
FT   DOMAIN     114    166       LAMININ EGF-LIKE 2.
FT   DOMAIN     167    185       LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN     202    793       DOMAIN II AND I (HEPTAT REPEATS).
FT   DOMAIN     794   1713       5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT   DOMAIN     794    970       LAMININ G-LIKE 1.
FT   DOMAIN     971   1139       LAMININ G-LIKE 2.
FT   DOMAIN    1140   1353       LAMININ G-LIKE 3.
FT   DOMAIN    1354   1529       LAMININ G-LIKE 4.
FT   DOMAIN    1530   1713       LAMININ G-LIKE 5.
FT   DOMAIN     231    327       COILED COIL (POTENTIAL).
FT   DOMAIN     396    548       COILED COIL (POTENTIAL).
FT   DOMAIN     594    621       COILED COIL (POTENTIAL).
FT   DOMAIN     702    765       COILED COIL (POTENTIAL).
FT   DOMAIN    1686   1713       COILED COIL (POTENTIAL).
FT   DISULFID    67     76       BY SIMILARITY.
FT   DISULFID    69     83       BY SIMILARITY.
FT   DISULFID    86     95       BY SIMILARITY.
FT   DISULFID    98    111       BY SIMILARITY.
FT   DISULFID   114    126       BY SIMILARITY.
FT   DISULFID   116    135       BY SIMILARITY.
FT   DISULFID   137    146       BY SIMILARITY.
FT   DISULFID   149    164       BY SIMILARITY.
FT   DISULFID   202    202       INTERCHAIN (PROBABLE).
FT   DISULFID   205    205       INTERCHAIN (PROBABLE).
FT   SITE       658    660       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD   542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   645    645       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   745    745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   882    882       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   964    964       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1108   1108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1131   1131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1325   1325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1477   1477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1667   1667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1     45       MGWLWIFGAALGQCLGYSSQQQRVPFLQPPGQSQLQASYVE
FT                               FRPS -> KVSSYGGYLTYQAKSFALPGDMVLLEKKPDVQL
FT                               TGQHMSIIYEETNTPRPDRLHHGRVHVVEGNFRHASSRAPV
FT                               SREELMTVLSRLADVRIQGLYFTETQRLTLSEVGLEEASDT
FT                               GSGRIALAVEICACPPAYAGDSC (IN ISOFORM B).
FT   CONFLICT      5      5       W -> R (IN REF. 2).
FT   CONFLICT    123    125       ATG -> GMC (IN REF. 2).
FT   CONFLICT    481    481       M -> K (IN REF. 2).
FT   CONFLICT    754    754       R -> L (IN REF. 2).
FT   CONFLICT    969    969       E -> Q (IN REF. 2).
FT   CONFLICT   1052   1052       D -> A (IN REF. 2).
FT   CONFLICT   1184   1184       G -> A (IN REF. 2).
SQ   SEQUENCE   1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;
```

```
  Query Match             18.1%;  Score 64.5;  DB 1;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 29;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFAFAALNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


Search completed: February  3, 2001, 02:24:28
Job time: 640 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:39 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	340	95.2	352	1	QQCVL1	AL1 protein - toma
2	240	67.2	361	1	QQCVPT	AL1 protein - pota
3	227	63.6	358	2	S07594	hypothetical prote
4	220	61.6	362	1	JQ1887	AL1 protein - toma
5	217	60.8	359	2	S39211	gene C1 protein -
6	212	59.4	351	2	JQ2327	AL1 protein - Indi
7	210	58.8	359	2	S22593	hypothetical prote
8	209	58.5	349	2	JQ2300	replicase - pepper
9	209	58.5	349	2	S31875	AL1 protein - pepp
10	204	57.1	359	2	S39235	gene C1 protein -
11	203	56.9	355	1	QQCVW1	AV1 protein - abut
12	200	56.0	358	1	JQ1870	AL1 protein - toma
13	200	56.0	385	2	S28360	AL1 protein - beet
14	193	54.1	360	2	S59885	replication-associ
15	191	53.5	357	1	QQCVC1	AL1 protein - toma
16	118	33.1	131	2	S45059	AC1 protein (clone
17	104	29.1	347	1	QQCVS1	AL1 protein - squa
18	67	18.8	447	2	T12544	hypothetical prote
19	64.5	18.1	587	2	JC1419	Fc gamma (IgG) rec
20	64.5	18.1	1713	2	A55347	adhesive ligand ep
21	63.5	17.8	1610	2	A46227	voltage-dependent
22	63.5	17.8	1646	2	JH0422	voltage-dependent
23	63.5	17.8	2161	2	JH0564	calcium channel al
24	63.5	17.8	2181	2	A38198	calcium channel al
25	63.5	17.8	2203	2	T42742	voltage-dependent
26	63	17.6	316	2	C82085	conserved hypothet
27	61.5	17.2	392	2	T45290	acetyl-CoA C-acety
28	60.5	16.9	840	2	T36175	probable large ATP
29	60	16.8	269	2	S61633	YNT20 protein - ye
30	60	16.8	295	2	D42452	C1 protein - tobac
31	60	16.8	355	2	T50746	phytoene synthase
32	60	16.8	355	2	S49621	phytoene synthase
33	60	16.8	1888	2	T14273	zinc finger protei
34	59.5	16.7	299	2	B71967	probable peptidyl-
35	59.5	16.7	470	2	T46814	gamma-aminobutyrat
36	59.5	16.7	2137	2	T05244	hypothetical prote
37	59.5	16.7	4385	2	T29042	hypothetical prote
38	58.5	16.4	354	2	A75087	acetyl ornithine d
39	58.5	16.4	647	2	T26240	hypothetical prote
40	58.5	16.4	706	2	H71707	hypothetical prote
41	58.5	16.4	1203	2	T17415	mycelial surface a
42	58.5	16.4	1232	2	S40766	hypothetical prote
43	58.5	16.4	1743	2	T15893	hypothetical prote
44	58	16.2	442	2	C45561	merozoite surface
45	58	16.2	442	2	B45561	merozoite surface

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.2%; Score 340; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 3.8e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||| |  |||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             67.2%;  Score 240;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 9.6e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: || ||:||||||| || |||||||||: :|| |::||:||:|||:|| :  |: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.6%;  Score 227;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 2.9e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: ::  |  |||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.6%;  Score 220;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 1.8e-17;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |:  |  |||||

Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||      .          ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.8%;  Score 217;  DB 2;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 4e-17;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |  :|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179


RESULT   6
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.4%;  Score 212;  DB 2;  Length 351;
  Best Local Similarity   62.7%;  Pred. No. 1.4e-16;
  Matches   42;  Conservative    8;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          || ||:|||||||| |::||| | |||: || |||:|:||  |  ||  |  ::||||||
Db    113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy     64 FDKTPEP 70
          | | | |
Db    173 FTKPPPP 179


RESULT   7
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.8%;  Score 210;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.5e-16;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
            | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   8
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante, R.F
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with bipa
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.5%;  Score 209;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 3.2e-16;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||:|: |: :   |  : ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   9
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.5%;  Score 209;  DB 2;  Length 349;
  Best Local Similarity   57.1%;  Pred. No. 3.2e-16;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: |||||:|||||||| |::||  |:|||::| |||||||:|: |: :   |  : ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT  10
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             57.1%;  Score 204;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.2e-15;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
            | || ||:|||||||| ||:||| |:|:|| || ||| :|::  |  |:  |  ::||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT  11
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.9%;  Score 203;  DB 1;  Length 355;
  Best Local Similarity   55.7%;  Pred. No. 1.5e-15;
  Matches   39;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
```

```
          |  |||||:|||||||| ||:||: |:|||||   : || |::|: |: |:     : |||
Db    110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179


RESULT  12
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.0%;  Score 200;  DB 1;  Length 358;
  Best Local Similarity   52.9%;  Pred. No. 3.4e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: ::     : |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176


RESULT  13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.0%;  Score 200;  DB 2;  Length 385;
  Best Local Similarity   52.9%;  Pred. No. 3.7e-15;
  Matches   37;  Conservative   16;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
          |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |: |       | :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             54.1%;  Score 193;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 2.2e-14;
  Matches   38;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          :| ||:|||||||| |::||| |||:|: ||  || |:||| |: ::|  |  ||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178


RESULT  15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             53.5%;  Score 191;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 3.6e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRI 63
          :|  |:|||||||| |::||| |||||: || ||| |:::|| |: |:|  |  |:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


Search completed: February  3, 2001, 02:17:40
Job time: 2212 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:45 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	210	58.8	359	3	US-08-809-103B-2	Sequence 2, Appli
2	210	58.8	359	3	US-08-809-103B-4	Sequence 4, Appli
3	210	58.8	359	3	US-08-809-103B-6	Sequence 6, Appli
4	210	58.8	359	3	US-08-809-103B-8	Sequence 8, Appli
5	64.5	18.1	1713	3	US-08-600-982-24	Sequence 24, Appl
6	64.5	18.1	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
7	63.5	17.8	2161	1	US-07-745-206A-2	Sequence 2, Appli
8	63.5	17.8	2161	1	US-08-455-543A-49	Sequence 49, Appl
9	63.5	17.8	2161	1	US-08-455-543A-51	Sequence 51, Appl
10	63.5	17.8	2161	2	US-08-223-305C-49	Sequence 49, Appl
11	63.5	17.8	2161	2	US-08-223-305C-51	Sequence 51, Appl
12	63.5	17.8	2161	2	US-08-311-363-2	Sequence 2, Appli
13	58	16.2	446	2	US-08-672-814D-11	Sequence 11, Appl
14	55	15.4	498	2	US-08-511-485-13	Sequence 13, Appl
15	54.5	15.3	335	2	US-09-014-969-15	Sequence 15, Appl
16	54	15.1	633	1	US-08-458-477A-5	Sequence 5, Appli
17	54	15.1	633	2	US-09-033-153-5	Sequence 5, Appli
18	54	15.1	765	2	US-08-663-112-2	Sequence 2, Appli
19	53.5	15.0	217	1	US-08-621-081A-23	Sequence 23, Appl
20	53	14.8	266	2	US-07-857-224B-40	Sequence 40, Appl
21	53	14.8	374	2	US-08-928-692-51	Sequence 51, Appl
22	53	14.8	454	3	US-08-446-100-30	Sequence 30, Appl
23	52.5	14.7	217	1	US-08-621-081A-22	Sequence 22, Appl
24	52.5	14.7	382	2	US-08-477-451-28	Sequence 28, Appl
25	52.5	14.7	1781	2	US-08-477-451-11	Sequence 11, Appl
26	52	14.6	439	1	US-07-762-681A-1	Sequence 1, Appli
27	52	14.6	454	3	US-08-446-100-26	Sequence 26, Appl
28	52	14.6	454	3	US-08-446-100-27	Sequence 27, Appl
29	52	14.6	454	3	US-08-446-100-28	Sequence 28, Appl
30	52	14.6	454	3	US-08-446-100-29	Sequence 29, Appl
31	52	14.6	454	3	US-08-446-100-31	Sequence 31, Appl
32	52	14.6	482	1	US-07-792-885A-1	Sequence 1, Appli
33	52	14.6	482	1	US-08-142-439A-7	Sequence 7, Appli
34	52	14.6	482	2	US-08-869-477-7	Sequence 7, Appli
35	52	14.6	933	3	US-08-293-728-2	Sequence 2, Appli
36	52	14.6	4928	3	US-09-036-987A-5	Sequence 5, Appli
37	51.5	14.4	373	3	US-08-258-287B-42	Sequence 42, Appl
38	51.5	14.4	373	3	US-08-368-704C-42	Sequence 42, Appl
39	51.5	14.4	649	3	US-08-911-853-15	Sequence 15, Appl
40	51.5	14.4	894	1	US-08-117-362-4	Sequence 4, Appli
41	51.5	14.4	894	1	US-08-486-924-4	Sequence 4, Appli
42	51	14.3	248	1	US-07-901-707-5	Sequence 5, Appli
43	51	14.3	248	1	US-07-988-430-5	Sequence 5, Appli
44	51	14.3	248	1	US-08-425-336-5	Sequence 5, Appli
45	51	14.3	248	1	US-08-378-761A-75	Sequence 75, Appl

ALIGNMENTS

```
RESULT   1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match            58.8%;  Score 210;  DB 3;  Length 359;
  Best Local Similarity  56.5%;  Pred. No. 2.3e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;
```

```
Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             58.8%;  Score 210;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.3e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             58.8%;  Score 210;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.3e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             58.8%;  Score 210;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.3e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             18.1%;  Score 64.5;  DB 3;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 7.6;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFAFAALNSNLDRIFD 65
             ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   6
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24
```

```
  Query Match             18.1%;  Score 64.5;  DB 4;  Length 1713;
  Best Local Similarity   34.4%;  Pred. No. 7.6;
  Matches   21;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       9 VDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFAFAALNSNLDRIFD 65
           ||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |:: :
Db     553 VDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLN 608

Qy      66 K 66
           :
Db     609 E 609


RESULT   7
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             17.8%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 14;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |    ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   8
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             17.8%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 14;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;
```

```
Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  9
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match             17.8%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 14;
  Matches   21;  Conservative   5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  10
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
```

```
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             17.8%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 14;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  11
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             17.8%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 14;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  12
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,363
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-51506
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             17.8%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 14;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  13
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
;  GENERAL INFORMATION:
;    APPLICANT:  Leung, David W.
;    APPLICANT:  Weeks, Reitha
;    TITLE OF INVENTION:  MAMMALIAN
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cell Therapeutics, Inc.
;      STREET:  201 Elliott Avenue West
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98119
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
;      COMPUTER:  PC Clone (486 microprocessor)
;      OPERATING SYSTEM:  MS-DOS Version 6.1, Windows NT
;      SOFTWARE:  WORD 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,814D
;      FILING DATE:  28-Jun-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Faciszewski, Stephen
;      REGISTRATION NUMBER:  36,131
;      REFERENCE/DOCKET NUMBER:  1803
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)282-7100
;      TELEFAX:  (206)284-6206
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  446
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:  CDP-diacylglycerol synthase (CDS)
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-672-814D-11


  Query Match             16.2%;  Score 58;  DB 2;  Length 446;
  Best Local Similarity   35.3%;  Pred. No. 9.5;
  Matches   18;  Conservative   10;  Mismatches   19;  Indels    4;  Gaps    2;

Qy      20 TSNDAAAEALNASSKEEALQIIREKIP-EKYLFAFAALNSNLDRIFDKTPE 69
           : :||| :  :|:   : :    |||| ||::   |    || :  |||||
Db      20 SGSDAANKRNSAADSSDHVDSEEEKIPEEKFVDELA---KNLPQGTDKTPE 67


RESULT  14
US-08-511-485-13
; Sequence 13, Application US/08511485
; Patent No. 5919912
;  GENERAL INFORMATION:
;    APPLICANT:  Korneluk, Robert G.
;    APPLICANT:  Mackenzie, Alexander E.
;    APPLICANT:  Baird, Stephen
;    TITLE OF INVENTION:  MAMMALIAN IAP GENE FAMILY, PRIMERS,
;    TITLE OF INVENTION:  PROBES, AND DETECTION METHODS
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/511,485
;      FILING DATE:  04-AUG-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  07540/002001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617/542-5070
;      TELEFAX:  617/542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  498 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  both
;    MOLECULE TYPE:  protein
US-08-511-485-13


  Query Match             15.4%;  Score 55;  DB 2;  Length 498;
  Best Local Similarity   26.8%;  Pred. No. 28;
  Matches   26;  Conservative    8;  Mismatches   23;  Indels   40;  Gaps    4;

Qy      14 ARGGCQTSNDAAAEALNASSK----------------EEALQ------IIREKIPE 47
           |:|    |   |  | ||||:                  :|||       ::|  |
Db     282 AKGPAYVSEVLATTAANASSQPATAPAPTLQADVLMDEAPAKEALTLGIDGGVVRNAIQR 341

Qy      48 KYL---FAFAALNSNLDRIFD----------KTPEP 70
```

```
          | |    ||: |:  |  |||           : |||
Db    342 KLLSSGCAFSTLDELLHDIFDDAGAGAALEVREPPEP 378


RESULT 15
US-09-014-969-15
; Sequence 15, Application US/09014969
; Patent No. 5965397
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John M.
;    APPLICANT:  LaVallie, Edward R.
;    APPLICANT:  Racie, Lisa A.
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    APPLICANT:  Agostino, Michael J.
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/014,969
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sprunger, Suzanne A.
;      REGISTRATION NUMBER:  41,323
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8284
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  335 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-014-969-15


  Query Match             15.3%;  Score 54.5;  DB 2;  Length 335;
  Best Local Similarity   34.0%;  Pred. No. 19;
  Matches   16;  Conservative    7;  Mismatches   19;  Indels    5;  Gaps    1;

Qy      24 AAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRIFDKTPEP 70
           |:   :|     ||||: || ||   ::|      |::||  |   |
Db     270 ASLPPVNMYDLFEALQVHREVIPTHTVYAL-----NIERIITKLWHP 311


Search completed: February  3, 2001, 02:19:47
Job time: 1709 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:28 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-6
Perfect score: 357
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FAFAALNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	210	58.8	359	17	R88870	Sardinian tomato y
2	210	58.8	359	17	R88871	Sardinian tomato y
3	210	58.8	359	17	R88872	Sardinian tomato y
4	206	57.7	353	18	W34338	Bean golden mosaic
5	206	57.7	353	18	W34332	Bean golden mosaic
6	206	57.7	353	18	W34333	Bean golden mosaic
7	206	57.7	353	18	W34334	Bean golden mosaic
8	206	57.7	353	18	W34335	Bean golden mosaic
9	204	57.1	353	8	P70407	ORF 4 gene product
10	204	57.1	361	18	W34336	Tomato mottle viru
11	204	57.1	361	18	W34324	Tomato mottle viru
12	204	57.1	361	18	W34325	Tomato mottle viru
13	204	57.1	361	18	W34326	Tomato mottle viru
14	197.5	55.3	361	8	P70562	Product of ORF 4 f
15	192	53.8	362	19	W56495	Tobacco leaf curli
16	191	53.5	357	18	W34329	Tomato yellow leaf
17	191	53.5	357	18	W34330	Tomato yellow leaf
18	191	53.5	357	18	W34331	Tomato yellow leaf
19	183	51.3	357	18	W34337	Tomato yellow leaf
20	67	18.8	447	21	Y87089	Human secreted pro
21	67	18.8	447	21	Y66744	Membrane-bound pro
22	67	18.8	447	21	Y50947	Human adult aorta
23	67	18.8	456	21	Y87190	Human secreted pro
24	64.5	18.1	1713	16	R70148	Deduced sequence o
25	63.5	17.8	619	13	R27651	Human calcium chan
26	63.5	17.8	2161	14	R33545	Sequence of the al
27	63.5	17.8	2161	16	R71001	Human neuronal cal
28	63.5	17.8	2161	16	R71002	Human neuronal cal
29	63.5	17.8	2161	19	W63137	Human calcium chan
30	63.5	17.8	2161	19	W63149	Human calcium chan
31	63	17.6	292	21	Y92317	Geminivirus Rep C1
32	63	17.6	335	21	Y92318	Mastrevirus Rep in
33	60.5	16.9	131	18	W34327	Tomato mottle viru
34	59.5	16.7	299	18	W55450	H. pylori ORF 02ae
35	59	16.5	318	20	Y31895	Soybean cyclin del
36	58.5	16.4	1203	21	Y83275	Candida albicans C
37	58.5	16.4	1203	21	Y84814	Amino acid sequenc
38	57	16.0	62	20	Y11882	Human 5' EST secre
39	56	15.7	767	20	Y06425	Mouse topoisomeras
40	55.5	15.5	204	20	W89814	Protein encoded by
41	55.5	15.5	266	19	W72937	Mycobacterium tube
42	55.5	15.5	266	20	Y21957	Amino acid sequenc
43	55.5	15.5	299	20	W89829	Protein encoded by
44	55.5	15.5	299	20	W89849	Antigen from clust
45	55.5	15.5	374	15	R61235	Site specific inte

ALIGNMENTS

RESULT 1
R88870
ID R88870 standard; Protein; 359 AA.
XX
AC R88870;
XX
DT 07-NOV-1996 (first entry)
XX
DE Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW modification; mutation; viral replication; deficient; inhibition;
KW viral resistance; geminivirus; tomato yellow leaf curl virus;
KW Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW AL1 protein; dominant negative phenotype.
XX
OS Sardinian tomato yellow leaf curl virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 227
FT /note= "wild-type Lys has been replaced by Ala"
XX
PN WO9608573-A1.
XX
PD 21-MAR-1996.
XX
PF 15-SEP-1995; 95WO-FR01192.
XX
PR 15-SEP-1994; 94FR-0011040.
XX
PA (CNRS) CENT NAT RECH SCI.
XX
PI Gronenborn B;
XX
DR WPI; 1996-179947/18.

```
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             58.8%;  Score 210;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1e-18;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 kvfqvppap 179


RESULT   2
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             58.8%;  Score 210;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1e-18;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
             ::|   | |
Db     171 kvfqvppap 179


RESULT   3
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
```

```
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             58.8%;  Score 210;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1e-18;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | || ||:||||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |  :|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   4
W34338
ID   W34338 standard; Protein; 353 AA.
XX
AC   W34338;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 protein.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93314.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             57.7%;  Score 206;  DB 18;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-18;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   5
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;
```

```
  Query Match             57.7%;  Score 206;  DB 18;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-18;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             57.7%;  Score 206;  DB 18;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-18;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   7
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             57.7%;  Score 206;  DB 18;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-18;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179
```

```
RESULT   8
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             57.7%;  Score 206;  DB 18;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 3.3e-18;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   9
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             57.1%;  Score 204;  DB 8;  Length 353;
  Best Local Similarity   57.1%;  Pred. No. 5.9e-18;
  Matches   40;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |:     : |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             57.1%;  Score 204;  DB 18;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 6.1e-18;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: ::     : |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             57.1%;  Score 204;  DB 18;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 6.1e-18;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: ::     : |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             57.1%;  Score 204;  DB 18;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 6.1e-18;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: ::     : |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             57.1%;  Score 204;  DB 18;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 6.1e-18;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: ::     : |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             55.3%;  Score 197.5;  DB 8;  Length 361;
  Best Local Similarity   63.1%;  Pred. No. 4.1e-17;
  Matches   41;  Conservative    9;  Mismatches   14;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLDRIF 64
           | |||||||||||  |::||| |||||: || :|| |:||| |::|:  |  || || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             53.8%;  Score 192;  DB 19;  Length 362;
  Best Local Similarity   50.0%;  Pred. No. 2e-16;
  Matches   42;  Conservative   10;  Mismatches   16;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFAFAALNSNLD 61
           | :| |||||||||  ||::||| |||:|: ||  || |:||| |: ::  |  ||||||
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 RI----------------FDKTPE 69
           ||                ||: ||
Db     171 rifapplevfvcpfssssfdqvpe 194


Search completed: February  3, 2001, 02:15:28
Job time: 3435 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:28 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	213	59.3	353	18	W34338	Bean golden mosaic
2	213	59.3	353	18	W34332	Bean golden mosaic
3	213	59.3	353	18	W34333	Bean golden mosaic
4	213	59.3	353	18	W34334	Bean golden mosaic
5	213	59.3	353	18	W34335	Bean golden mosaic
6	211	58.8	353	8	P70407	ORF 4 gene product
7	211	58.8	359	17	R88870	Sardinian tomato y
8	211	58.8	359	17	R88871	Sardinian tomato y
9	211	58.8	359	17	R88872	Sardinian tomato y
10	211	58.8	361	18	W34336	Tomato mottle viru
11	211	58.8	361	18	W34324	Tomato mottle viru
12	211	58.8	361	18	W34325	Tomato mottle viru
13	211	58.8	361	18	W34326	Tomato mottle viru
14	208.5	58.1	361	8	P70562	Product of ORF 4 f
15	194	54.0	362	19	W56495	Tobacco leaf curli
16	193	53.8	357	18	W34329	Tomato yellow leaf
17	193	53.8	357	18	W34330	Tomato yellow leaf
18	193	53.8	357	18	W34331	Tomato yellow leaf
19	185	51.5	357	18	W34337	Tomato yellow leaf
20	68	18.9	447	21	Y87089	Human secreted pro
21	68	18.9	447	21	Y66744	Membrane-bound pro
22	68	18.9	447	21	Y50947	Human adult aorta
23	68	18.9	456	21	Y87190	Human secreted pro
24	63.5	17.7	619	13	R27651	Human calcium chan
25	63.5	17.7	2161	14	R33545	Sequence of the al
26	63.5	17.7	2161	16	R71001	Human neuronal cal
27	63.5	17.7	2161	16	R71002	Human neuronal cal
28	63.5	17.7	2161	19	W63137	Human calcium chan
29	63.5	17.7	2161	19	W63149	Human calcium chan
30	62.5	17.4	512	19	W68473	HIV-1 strain YBF30
31	60.5	16.9	131	18	W34327	Tomato mottle viru
32	59	16.4	292	21	Y92317	Geminivirus Rep C1
33	59	16.4	335	21	Y92318	Mastrevirus Rep in
34	58.5	16.3	299	18	W55450	H. pylori ORF 02ae
35	58	16.2	374	19	W37994	Mutant Aspergillus
36	58	16.2	374	20	Y39875	A. oryzae P4-8.1 l
37	57	15.9	108	16	R82847	Murine mutant cyto
38	57	15.9	454	21	Y49496	Human wild-type ca
39	57	15.9	454	21	Y49497	Human carboxyleste
40	57	15.9	454	21	Y49498	Human carboxyleste
41	57	15.9	454	21	Y49499	Human carboxyleste
42	57	15.9	454	21	Y49500	Human carboxyleste
43	56.5	15.7	188	21	Y95043	Candida albicans p
44	56	15.6	108	16	R82846	Murine mutant cyto
45	56	15.6	213	17	W00648	Human G-CSF recept

ALIGNMENTS

RESULT 1
W34338
ID W34338 standard; Protein; 353 AA.
XX
AC W34338;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 protein.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93314.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

```
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             59.3%;  Score 213;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.5e-19;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 AAIFDKTPEP 70
            || | |||
Db    170 erifvkvpep 179


RESULT   2
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             59.3%;  Score 213;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.5e-19;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 AAIFDKTPEP 70
            || | |||
Db    170 erifvkvpep 179


RESULT   3
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
```

```
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match           59.3%;  Score 213;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.5e-19;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 AAIFDKTPEP 70
            || | |||
Db    170 erifvkvpep 179


RESULT   4
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match           59.3%;  Score 213;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.5e-19;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 AAIFDKTPEP 70
            || | |||
Db    170 erifvkvpep 179


RESULT   5
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match           59.3%;  Score 213;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.5e-19;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169
```

```
Qy      61 AAIFDKTPEP 70
           || | |||
Db     170 erifvkvpep 179


RESULT   6
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             58.8%;  Score 211;  DB 8;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 2.7e-19;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |: |: | ||: | |
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 AAIFDKTPEP 70
           || | |||
Db     170 eriffkvpep 179


RESULT   7
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             58.8%;  Score 211;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  | |:  |||:|| |
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 kvfqvppap 179


RESULT   8
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
```

```
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             58.8%;  Score 211;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 kvfqvppap 179


RESULT   9
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             58.8%;  Score 211;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 kvfqvppap 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             58.8%;  Score 211;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: | |
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 AAIFDKTPEP 70
            || | |||
Db    170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             58.8%;  Score 211;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: | |
Db    110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy     61 AAIFDKTPEP 70
            || | |||
Db    170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             58.8%;  Score 211;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: | |
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             58.8%;  Score 211;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 2.8e-19;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||: |:|||||| : || ::||: |: :: | ||: | |
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             58.1%;  Score 208.5;  DB 8;  Length 361;
  Best Local Similarity   64.6%;  Pred. No. 5.9e-19;
  Matches   42;  Conservative   10;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAIF 64
           | |||||||||||| |::||| |||||: || :|| |:||| |::|: |||||| |: ||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             54.0%;  Score 194;  DB 19;  Length 362;
  Best Local Similarity   50.0%;  Pred. No. 4.5e-17;
  Matches   42;  Conservative   10;  Mismatches   16;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | :| |||||||||  ||::||| |||:|: ||  || |:||| |: :: ||||||| |
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 AI----------------FDKTPE 69
           |                 ||: ||
Db     171 rifapplevfvcpfssssfdqvpe 194


Search completed: February  3, 2001, 02:15:29
Job time: 3436 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:52 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	281	77.2	226	12	O09727	O09727 leonurus mo
2	281	77.2	226	12	Q9WHF6	Q9whf6 tomato mild
3	281	77.2	361	12	Q67574	Q67574 bean golden
4	279	76.6	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	277	76.1	364	12	Q9QS55	Q9qs55 sweet potat
6	274	75.3	185	12	Q98693	Q98693 sida golden
7	266	73.1	149	12	P88975	P88975 macroptiliu
8	266	73.1	233	12	Q9YLA4	Q9yla4 macroptiliu
9	261	71.7	190	12	Q9ZO89	Q9z089 tobacco lea
10	261	71.7	190	12	Q9ZO84	Q9z084 tobacco lea
11	260	71.4	190	12	Q9W827	Q9w827 tobacco lea
12	260	71.4	208	12	Q9Z0C4	Q9z0c4 tobacco lea
13	260	71.4	234	12	O39180	O39180 potato yell
14	258	70.9	208	12	Q9Z0C0	Q9z0c0 tobacco lea
15	258	70.9	208	12	Q9Z0B8	Q9z0b8 tobacco lea
16	256	70.3	208	12	Q9Z0B6	Q9z0b6 tobacco lea
17	255	70.1	203	12	Q9ZO83	Q9z083 tobacco lea
18	251	69.0	363	12	O73577	O73577 cotton leaf
19	250	68.7	363	12	O72705	O72705 cotton leaf
20	250	68.7	363	12	O72719	O72719 cotton leaf
21	249	68.4	349	12	Q88888	Q88888 tomato pseu
22	248	68.1	208	12	Q9Z0A0	Q9z0a0 tobacco lea
23	247	67.9	190	12	Q9ZO86	Q9z086 tobacco lea
24	247	67.9	208	12	Q9Z0C6	Q9z0c6 tobacco lea
25	247	67.9	359	12	Q9YZV4	Q9yzv4 tomato yell
26	247	67.9	359	12	Q9YZV2	Q9yzv2 tomato yell
27	247	67.9	359	12	Q9YUX7	Q9yux7 tomato yell
28	247	67.9	359	12	Q9YL27	Q9yl27 tomato yell
29	246	67.6	359	12	Q88942	Q88942 tomato yell
30	245	67.3	362	12	Q9IN48	Q9in48 okra enatio
31	244	67.0	307	12	Q9IET7	Q9iet7 cotton leaf
32	244	67.0	361	12	O72723	O72723 cotton leaf
33	243	66.8	190	12	Q9Z0A7	Q9z0a7 tobacco lea
34	243	66.8	354	12	Q9IN42	Q9in42 south afric
35	242	66.5	358	12	Q9JEA2	Q9jea2 cassava gem
36	242	66.5	362	12	O56816	O56816 chayote mos
37	241	66.2	231	12	Q96620	Q96620 african tom
38	240	65.9	353	12	O72692	O72692 beet curly
39	240	65.9	354	12	Q65438	Q65438 beet curly
40	239	65.7	153	12	O41475	O41475 wissadula g
41	238	65.4	360	12	Q9YNW5	Q9ynw5 sida golden
42	238	65.4	362	12	Q9YYT7	Q9yyt7 althea rose
43	238	65.4	363	12	O73494	O73494 okra yellow
44	236	64.8	306	12	Q9IET1	Q9iet1 cotton leaf
45	236	64.8	351	12	Q9Q9R3	Q9q9r3 dicliptera

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 77.2%; Score 281; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 6.3e-23;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||||||:||| || ||   |||||  |  |||||:||:||||||||||||||||
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 6.3e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| ||:||   ||||||||||||::|||:|:|:::|||:|||:|||
Db    111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.2%;  Score 281;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 1.1e-22;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          || |||||||||| ||:||   ||||||||||||:|||:|||:|||:|||:||||||
Db    113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |:|
Db    173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             76.6%;  Score 279;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 1e-22;
  Matches   52;  Conservative    7;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          || ||:|||||||| || ||   ||||||||||||:|||:|||:|||:|||:||||||
Db    113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy     64 FDKTPEP 70
          | | |||
Db    173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             76.1%;  Score 277;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 2.9e-22;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||||||||||| ||:||   ||||| ||| ||||||||:|||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             75.3%;  Score 274;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 2.9e-22;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| ||:||   ||||: :||:||:|||||:||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             73.1%;  Score 266;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 1.7e-21;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| |||||   ||||: :|| |::|::||:|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             73.1%;  Score 266;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 2.8e-21;
```

```
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| |||||   ||||: :|| |::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   9
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             71.7%;  Score 261;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 7.7e-21;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|||||||||| :||   |||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  10
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             71.7%;  Score 261;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 7.7e-21;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || ||:|||||||||| :||   |||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             71.4%;  Score 260;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 9.9e-21;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || || |||||||||||||| :||   |||||:|| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  12
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
```

```
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             71.4%;  Score 260;  DB 12;  Length 208;
  Best Local Similarity   60.0%;  Pred. No. 1.1e-20;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || |||||||||||| :||   |||||:|| ||| |||||:|: ::|||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  13
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             71.4%;  Score 260;  DB 12;  Length 234;
  Best Local Similarity   67.1%;  Pred. No. 1.2e-20;
  Matches   47;  Conservative   13;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:||||||| ||:||   |||| :||||::||:||:|||:|||:|||:|:|
Db     110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy      61 DRIFDKTPEP 70
           |||| | | |
Db     170 DRIFMKAPNP 179


RESULT  14
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             70.9%;  Score 258;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 1.8e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||   |||||||| :|| |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;
```

Query Match 70.9%; Score 258; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 1.8e-20;
Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || || ||:|||||||||| :||   |||||||| :|| |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186
```

Search completed: February 3, 2001, 02:22:53
Job time: 1890 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:29 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	95.9	352	1	VAL1_TGMV	P03567 tomato gold
2	249	68.4	361	1	VAL1_PYMVV	P27258 potato yell
3	242	66.5	358	1	VAL1_CLVK	P14982 cassava lat
4	242	66.5	358	1	VAL1_CLVN	P14972 cassava lat
5	235	64.6	362	1	VAL1_TYLCA	P36279 tomato yell
6	228	62.6	349	1	VAL1_PHUV	Q06923 pepper huas
7	228	62.6	359	1	VAL1_TYLCU	P38609 tomato yell
8	221	60.7	353	1	VAL1_BGMV	P05175 bean golden
9	221	60.7	359	1	VAL1_TYLCM	P27260 tomato yell
10	220	60.4	355	1	VAL1_ABMVW	P21947 abutilon mo
11	217	59.6	358	1	VAL1_BCTV	P14991 beet curly
12	217	59.6	361	1	VAL1_TMOV	Q06657 tomato mott
13	206	56.6	357	1	VAL1_TYLCV	P27259 tomato yell
14	125	34.3	347	1	VAL1_SLCV	P29048 squash leaf
15	63	17.3	342	1	RSMC_ECOLI	P39406 escherichia
16	62	17.0	295	1	VAL1_TYDVA	P31617 tobacco yel
17	62	17.0	387	1	Y4PF_RHISN	P55615 rhizobium s
18	61	16.8	555	1	GLPD_BACSU	P18158 bacillus su
19	61	16.8	836	1	GCSR_HUMAN	Q99062 homo sapien
20	60	16.5	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
21	60	16.5	799	1	AFSK_STRCO	P54741 streptomyce
22	60	16.5	807	1	AFSK_STRGR	P54742 streptomyce
23	60	16.5	1117	1	TERT_TETTH	O77448 tetrahymena
24	59.5	16.3	79	1	AKHD_DROME	P17975 drosophila
25	58.5	16.1	323	1	VAL1_PASVK	Q00338 panicum str
26	58.5	16.1	397	1	METL_RAT	P13444 rattus norv
27	58.5	16.1	1610	1	CCAD_MESAU	Q99244 mesocricetu
28	58.5	16.1	2161	1	CCAD_HUMAN	Q01668 homo sapien
29	58.5	16.1	2203	1	CCAD_RAT	P27732 rattus norv
30	58	15.9	256	1	YKJ9_YEAST	P34247 saccharomyc
31	58	15.9	354	1	CARA_METJA	Q58425 methanococc
32	58	15.9	543	1	NFL_HUMAN	P07196 homo sapien
33	58	15.9	1287	1	SKI2_YEAST	P35207 saccharomyc
34	57	15.7	784	1	LON_RICPR	Q9zd92 rickettsia
35	57	15.7	970	1	DA81_YEAST	P21657 saccharomyc
36	57	15.7	1517	1	RPOC_CAMJE	Q9pi30 campylobact
37	56.5	15.5	266	1	ETFB_MYCLE	O33095 mycobacteri
38	56.5	15.5	266	1	ETFB_MYCTU	O53276 mycobacteri
39	56.5	15.5	395	1	METL_HUMAN	Q00266 homo sapien
40	56	15.4	247	1	YCP4_YEAST	P25349 saccharomyc
41	56	15.4	299	1	Y175_HELPY	P56112 helicobacte
42	56	15.4	336	1	Y625_METJA	Q58042 methanococc
43	56	15.4	524	1	MPPA_RAT	P20069 rattus norv
44	56	15.4	618	1	ORC2_DROME	Q24168 drosophila
45	56	15.4	798	1	VP16_YEAST	Q03308 saccharomyc

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.9%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 2.1e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             68.4%;  Score 249;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 2.4e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||:|||||||| || ||   ||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             66.5%;  Score 242;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.4e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT ' 4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             66.5%;  Score 242;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 1.4e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             64.6%;  Score 235;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 8.6e-19;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||:|||||||| |::||   :|||  || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             62.6%;  Score 228;  DB 1;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 4.9e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||:|||||||| |::||   :|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   7
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             62.6%;  Score 228;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 5.1e-18;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:|| :|:|| || ||| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|  | |
Db    171 RVFQVPPAP 179


RESULT   8
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.7%;  Score 221;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 3e-17;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:|||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db    110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFFKVPEP 179


RESULT   9
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             60.7%;  Score 221;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3e-17;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  ||||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|  | |
Db    171 KVFQVPPAP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             60.4%;  Score 220;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 3.8e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||   :||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.6%;  Score 217;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 8.3e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||:|||||||| ||:||   :||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             59.6%;  Score 217;  DB 1;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 8.4e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             56.6%;  Score 206;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 1.4e-15;
  Matches   39;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |:||||||||| |::||   ||||: || ||| |::|| |: |: ||||||:||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.3%;  Score 125;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 1.2e-06;
  Matches   25;  Conservative   15;  Mismatches   22;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::| |    |     :|:   |:|| |  ||| ||:   |: :: |:||| :|:::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
RSMC_ECOLI
ID   RSMC_ECOLI     STANDARD;      PRT;   342 AA.
AC   P39406;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RIBOSOMAL RNA SMALL SUBUNIT METHYLTRANSFERASE C (EC 2.1.1.52) (RRNA
DE   (GUANINE-N2-)-METHYLTRANSFERASE) (16S RRNA M2G1207 METHYLTRANSFERASE).
GN   RSMC.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=95334362; PubMed=7610040;
RA   Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA   Blattner F.R.;
RT   "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT   region from 92.8 through 100 minutes.";
RL   Nucleic Acids Res. 23:2105-2119(1995).
RN   [2]
RP   SEQUENCE OF 1-19, AND CHARACTERIZATION.
RA   Tscherne J.S., Nurse K., Popienick P., Ofengand J.;
RT   "Purification, cloning, and characterization of the 16S RNA m2G1207
RT   methyltransferase from Escherichia coli.";
RL   J. Biol. Chem. 274:924-929(1999).
CC   -!- FUNCTION: SPECIFICALLY METHYLATES THE GUANOSINE IN POSITION 1207
CC       OF 16S RRNA IN THE 30S PARTICLE.
CC   -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + RRNA = S-ADENOSYL-L-
CC       HOMOCYSTEINE + RRNA CONTAINING N2-METHYLGUANINE.
CC   -!- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.
CC   -!- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY. RSMC
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U14003; AAA97267.1; -.
DR   EMBL; AE000507; AAC77324.1; -.
DR   ECOGENE; EG12596; RSMC.
DR   PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW   rRNA processing; Transferase; Methyltransferase; Magnesium.
```

```
FT   INIT_MET      0      0
SQ   SEQUENCE   342 AA;  37493 MW;  C7A318155700302D CRC64;


  Query Match             17.3%;  Score 63;  DB 1;  Length 342;
  Best Local Similarity   31.1%;  Pred. No. 8.2;
  Matches   14;  Conservative    7;  Mismatches   16;  Indels    8;  Gaps    1;

Qy       4 WGEFQVDGRS--------ARGGCQTSNDLLLEALNASSKEEALQI 40
           |||: ||| :         :| |    : |||  |   :| : | :
Db     158 WGEYSVDGLTVKTLPGVFSRDGLDVGSQLLLSTLTPHTKGKVLDV 202


Search completed: February  3, 2001, 02:24:30
Job time: 642 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:41 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	349	95.9	352	1	QQCVL1	AL1 protein - toma
2	249	68.4	361	1	QQCVPT	AL1 protein - pota
3	242	66.5	358	2	S07594	hypothetical prote
4	235	64.6	362	1	JQ1887	AL1 protein - toma
5	228	62.6	349	2	JQ2300	replicase - pepper
6	228	62.6	349	2	S31875	AL1 protein - pepp
7	228	62.6	359	2	S39211	gene C1 protein -
8	221	60.7	359	2	S22593	hypothetical prote
9	220	60.4	355	1	QQCVW1	AV1 protein - abut
10	218	59.9	351	2	JQ2327	AL1 protein - Indi
11	217	59.6	358	1	JQ1870	AL1 protein - toma
12	217	59.6	385	2	S28360	AL1 protein - beet
13	215	59.1	359	2	S39235	gene C1 protein -
14	208	57.1	360	2	S59885	replication-associ
15	206	56.6	357	1	QQCVC1	AL1 protein - toma
16	138	37.9	131	2	S45059	AC1 protein (clone
17	125	34.3	347	1	QQCVS1	AL1 protein - squa
18	64	17.6	587	2	JC1419	Fc gamma (IgG) rec
19	63.5	17.4	1229	2	T48959	kinesin-like prote
20	63	17.3	343	2	S56595	hypothetical 37.6K
21	62	17.0	295	2	D42452	C1 protein - tobac
22	62	17.0	416	2	A82892	hypothetical prote
23	61.5	16.9	333	2	B42476	hypothetical prote
24	61.5	16.9	481	2	A70091	probable phosphoes
25	61.5	16.9	1279	2	E64709	type IIS restricti
26	61	16.8	555	2	C45868	glycerol-3-phospha
27	61	16.8	771	2	B38252	granulocyte colony
28	61	16.8	783	2	JH0329	granulocyte colony
29	61	16.8	863	2	C38252	granulocyte colony
30	60.5	16.6	447	2	T12544	hypothetical prote
31	60.5	16.6	1792	2	T13939	myosin V - fruit f
32	60	16.5	470	2	T46814	gamma-aminobutyrat
33	60	16.5	799	2	T48889	serine/threonine p
34	60	16.5	1117	2	T14891	telomerase (EC 2.7
35	60	16.5	1265	2	T02131	hypothetical prote
36	59.5	16.3	79	1	A58656	adipokinetic hormo
37	59.5	16.3	210	2	T05021	hypothetical prote
38	59.5	16.3	388	2	C69196	hypothetical prote
39	59.5	16.3	546	2	A69484	hypothetical prote
40	59	16.2	373	2	A71690	hypothetical prote
41	59	16.2	506	2	T19787	hypothetical prote
42	59	16.2	785	2	T18252	conserved hypothet
43	59	16.2	1766	2	S03701	141K protein - pea
44	58.5	16.1	323	2	JQ1552	C1 protein - Panic
45	58.5	16.1	397	2	S06114	methionine adenosy

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.9%; Score 349; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.8e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 68.4%; Score 249; DB 1; Length 361;
Best Local Similarity 66.7%; Pred. No. 2.8e-20;
Matches 46; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||:|||||||| || ||   |||| :|| |::||:||:|||:|||:||| ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178
```

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 66.5%; Score 242; DB 2; Length 358;
Best Local Similarity 61.4%; Pred. No. 1.7e-19;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||   :|||: || ||| :||| :|: :: ||||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178
```

RESULT 4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 64.6%; Score 235; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 1e-18;
Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || |||||:|||||||| |::||   :|||  || ||| ::||  |: |: |||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194
```

RESULT 5
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.6%; Score 228; DB 2; Length 349;
Best Local Similarity 60.0%; Pred. No. 6e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| |::||   :|||::| |||||||:|: |: |:  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179
```

RESULT 6
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 62.6%; Score 228; DB 2; Length 349;
Best Local Similarity 60.0%; Pred. No. 6e-18;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| |::||   :|||::| |||||||:|: |: |:  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179
```

RESULT 7
S39211
gene C1 protein - tomato yellow leaf curl virus

```
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from the
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.6%;  Score 228;  DB 2;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 6.2e-18;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||   :|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 221;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.8e-17;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as wel
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.4%;  Score 220;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 4.8e-17;
  Matches   41;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||:|||||||| ||:||   :||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  10
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-t
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.9%;  Score 218;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 7.9e-17;
  Matches   41;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||:|||||||| |::||    |||: || |||:|:||  |  || ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.6%;  Score 217;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 1.1e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;
```

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||:|||||||| |::||   :|||| | : || ::||: |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176
```

RESULT 12
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.6%;  Score 217;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.1e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||:|||||||| ||:||   :||||:| ::||||::|: |: | | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206
```

RESULT 13
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.1%;  Score 215;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 1.8e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||   :|:|| || ||| :|:: |  |:  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179
```

RESULT 14

S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.1%;  Score 208;  DB 2;  Length 360;
  Best Local Similarity   59.1%;  Pred. No. 1.1e-15;
  Matches   39;  Conservative   12;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| ||:|||||||| |::||   ||:|: ||  || |:||| |: :: |||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             56.6%;  Score 206;  DB 1;  Length 357;
  Best Local Similarity   63.9%;  Pred. No. 1.8e-15;
  Matches   39;  Conservative   10;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |:|||||||| |::||   |||||: || ||| |::|| |: |:|||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171
```

Search completed: February 3, 2001, 02:17:42
Job time: 2214 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:51 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	221	60.7	359	3	US-08-809-103B-2	Sequence 2, Appli
2	221	60.7	359	3	US-08-809-103B-4	Sequence 4, Appli
3	221	60.7	359	3	US-08-809-103B-6	Sequence 6, Appli
4	221	60.7	359	3	US-08-809-103B-8	Sequence 8, Appli
5	61	16.8	602	2	US-08-419-652-6	Sequence 6, Appli
6	61	16.8	771	1	US-07-923-976-6	Sequence 6, Appli
7	61	16.8	783	5	5422248-2	Patent No. 5422248
8	61	16.8	836	1	US-07-923-976-4	Sequence 4, Appli
9	61	16.8	863	1	US-07-923-976-8	Sequence 8, Appli
10	58.5	16.1	2161	1	US-07-745-206A-2	Sequence 2, Appli
11	58.5	16.1	2161	1	US-08-455-543A-49	Sequence 49, Appl
12	58.5	16.1	2161	1	US-08-455-543A-51	Sequence 51, Appl
13	58.5	16.1	2161	2	US-08-223-305C-49	Sequence 49, Appl
14	58.5	16.1	2161	2	US-08-223-305C-51	Sequence 51, Appl
15	58.5	16.1	2161	2	US-08-311-363-2	Sequence 2, Appli
16	56	15.4	323	1	US-07-667-276A-7	Sequence 7, Appli
17	56	15.4	844	3	US-09-029-267-20	Sequence 20, Appl
18	55	15.1	751	3	US-09-036-987A-24	Sequence 24, Appl
19	54	14.8	244	1	US-08-910-973-22	Sequence 22, Appl
20	54	14.8	432	1	US-08-167-919A-11	Sequence 11, Appl
21	54	14.8	432	3	US-08-715-106-11	Sequence 11, Appl
22	54	14.8	1713	3	US-08-600-982-24	Sequence 24, Appl
23	54	14.8	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
24	53.5	14.7	439	1	US-07-637-870-9	Sequence 9, Appli
25	53.5	14.7	439	1	US-07-637-399-6	Sequence 6, Appli
26	53.5	14.7	439	1	US-08-112-703-6	Sequence 6, Appli
27	53.5	14.7	3898	2	US-08-876-991-2	Sequence 2, Appli
28	53.5	14.7	3898	2	US-09-059-853-2	Sequence 2, Appli
29	53	14.6	126	1	US-08-083-949-1	Sequence 1, Appli
30	53	14.6	237	1	US-08-910-973-13	Sequence 13, Appl
31	53	14.6	255	1	US-07-901-707-6	Sequence 6, Appli
32	53	14.6	255	1	US-07-988-430-6	Sequence 6, Appli
33	53	14.6	255	1	US-08-425-336-6	Sequence 6, Appli
34	53	14.6	255	4	PCT-US92-09487-6	Sequence 6, Appli
35	53	14.6	783	5	5231168-2	Patent No. 5231168
36	52.5	14.4	770	1	US-08-445-135-2	Sequence 2, Appli
37	52.5	14.4	870	2	US-08-266-311-1	Sequence 1, Appli
38	52.5	14.4	870	2	US-08-467-527A-1	Sequence 1, Appli
39	52.5	14.4	3079	4	PCT-US94-00198-4	Sequence 4, Appli
40	52	14.3	271	1	US-08-276-919-10	Sequence 10, Appl
41	52	14.3	271	1	US-08-776-088-13	Sequence 13, Appl
42	52	14.3	271	4	PCT-US95-09145A-13	Sequence 13, Appl
43	52	14.3	274	1	US-08-776-088-15	Sequence 15, Appl
44	52	14.3	274	4	PCT-US95-09145A-15	Sequence 15, Appl
45	52	14.3	325	1	US-08-276-919-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT   1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match             60.7%;  Score 221;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.2e-21;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             60.7%;  Score 221;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.2e-21;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             60.7%;  Score 221;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.2e-21;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             60.7%;  Score 221;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.2e-21;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || :||:|||||||| ||:|| :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
;  GENERAL INFORMATION:
;    APPLICANT:  Chua, Anne O
;    APPLICANT:  Gubler, Ulrich A
;    TITLE OF INVENTION:  INTERLEUKIN-12 RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07110-1199
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/419,652
;      FILING DATE:  11-APR-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/248,532
;      FILING DATE:  31-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/094,713
;      FILING DATE:  19-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kass, Alan P
;      REGISTRATION NUMBER:  32142
;      REFERENCE/DOCKET NUMBER:  CD 9174
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 235-4205
;      TELEFAX:  (201) 235-3500
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  602 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..602
;      OTHER INFORMATION:  /note= "Represents residues 98 to
;      OTHER INFORMATION:  731 of human granulocyte colony-stimulating
;      OTHER INFORMATION:  factor-receptor."
US-08-419-652-6


  Query Match             16.8%;  Score 61;  DB 2;  Length 602;
  Best Local Similarity   26.0%;  Pred. No. 7.4;
  Matches   13;  Conservative    9;  Mismatches   22;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           | :    :|| |||  | :|: :    :          || |:|  : |:
Db      37 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 80


RESULT   6
US-07-923-976-6
; Sequence 6, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
```

```
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-6


  Query Match             16.8%;  Score 61;  DB 1;  Length 771;
  Best Local Similarity   26.0%;  Pred. No. 10;
  Matches   13;  Conservative    9;  Mismatches   22;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           |  :    :|| |||  | :|: :   :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT   7
5422248-2
;Patent No. 5422248
;    APPLICANT: SMITH, CRAIG A.;LARSEN, ALF D.;SIMS, JOHN E.;
;BENSON, CURTIS M.
;    TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
;STIMULATING FACTOR RECEPTORS
;    NUMBER OF SEQUENCES: 6
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/6,183
;      FILING DATE: 15-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 587,329
;      FILING DATE: 24-SEP-1990
;      APPLICATION NUMBER: 522,952
;      FILING DATE: 03-APR-1990
;      APPLICATION NUMBER: 416,306
;      FILING DATE: 03-OCT-1989
;      APPLICATION NUMBER: 412,816
;      FILING DATE: 26-SEP-1989
;SEQ ID NO:2:
;      LENGTH: 783
5422248-2


  Query Match             16.8%;  Score 61;  DB 5;  Length 783;
  Best Local Similarity   26.0%;  Pred. No. 11;
  Matches   13;  Conservative    9;  Mismatches   22;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           |  :    :|| |||  | :|: :   :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT   8
US-07-923-976-4
; Sequence 4, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  836 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-4


  Query Match             16.8%;  Score 61;  DB 1;  Length 836;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   13;  Conservative    9;  Mismatches   22;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           |  :    :|| |||  | :|: :   :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT   9
US-07-923-976-8
; Sequence 8, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  863 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-8


  Query Match             16.8%;  Score 61;  DB 1;  Length 863;
  Best Local Similarity   26.0%;  Pred. No. 12;
  Matches   13;  Conservative    9;  Mismatches   22;  Indels    6;  Gaps    1;

Qy       7 FQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           |  :    :|| |||  | :|: :    :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT  10
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             16.1%;  Score 58.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 88;
  Matches   21;  Conservative    6;  Mismatches   18;  Indels   29;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTS--------------NDLLL--------------EALNASSK 34
           |     ||  | ||  :|              | :||             |:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  11
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             16.1%;  Score 58.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 88;
  Matches   21;  Conservative    6;  Mismatches   18;  Indels   29;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTS---------------NDLLL--------------EALNASSK 34
           |     ||  | ||  :|               | :||              |:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  12
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match             16.1%;  Score 58.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 88;
  Matches   21;  Conservative    6;  Mismatches   18;  Indels   29;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTS---------------NDLLL--------------EALNASSK 34
           |     ||  | ||  :|               | :||              |:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  13
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
```

```
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             16.1%;  Score 58.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 88;
  Matches   21;  Conservative    6;  Mismatches   18;  Indels   29;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTS---------------NDLLL--------------EALNASSK 34
           |     ||  | ||  :|                 | :||              |:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  14
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             16.1%;  Score 58.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 88;
  Matches   21;  Conservative    6;  Mismatches   18;  Indels   29;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTS---------------NDLLL--------------EALNASSK 34
           |     ||  | ||  :|                 | :||              |:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  15
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/311,363
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/745,206
;       FILING DATE:  15-AUG-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Seidman, Stephanie L.
;       REGISTRATION NUMBER:  33,779
;       REFERENCE/DOCKET NUMBER:  6362-51506
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619)238-0999
;       TELEFAX:  (619)238-0062
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2161 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             16.1%;  Score 58.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 88;
  Matches   21;  Conservative    6;  Mismatches   18;  Indels   29;  Gaps    2;

Qy       4 WGEFQVDGRSARGGCQTS----------------NDLLL--------------EALNASSK 34
           |     ||  | ||  :|                | :||              |:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


Search completed: February  3, 2001, 02:19:52
Job time: 1714 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:30 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-10
Perfect score: 364
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	223	61.3	353	18	W34338	Bean golden mosaic
2	223	61.3	353	18	W34332	Bean golden mosaic
3	223	61.3	353	18	W34333	Bean golden mosaic
4	223	61.3	353	18	W34334	Bean golden mosaic
5	223	61.3	353	18	W34335	Bean golden mosaic
6	221	60.7	353	8	P70407	ORF 4 gene product
7	221	60.7	359	17	R88870	Sardinian tomato y
8	221	60.7	359	17	R88871	Sardinian tomato y
9	221	60.7	359	17	R88872	Sardinian tomato y
10	221	60.7	361	18	W34336	Tomato mottle viru
11	221	60.7	361	18	W34324	Tomato mottle viru
12	221	60.7	361	18	W34325	Tomato mottle viru
13	221	60.7	361	18	W34326	Tomato mottle viru
14	212.5	58.4	361	8	P70562	Product of ORF 4 f
15	207	56.9	362	19	W56495	Tobacco leaf curli
16	206	56.6	357	18	W34329	Tomato yellow leaf
17	206	56.6	357	18	W34330	Tomato yellow leaf
18	206	56.6	357	18	W34331	Tomato yellow leaf
19	198	54.4	357	18	W34337	Tomato yellow leaf
20	65	17.9	292	21	Y92317	Geminivirus Rep C1
21	65	17.9	335	21	Y92318	Mastrevirus Rep in
22	63	17.3	134	16	R72393	Agmenellum quadrup
23	62	17.0	108	16	R82847	Murine mutant cyto
24	61.5	16.9	1279	19	W98761	H. pylori GHPO 109
25	61	16.8	108	16	R82846	Murine mutant cyto
26	61	16.8	213	17	W00648	Human G-CSF recept
27	61	16.8	333	17	R99141	Ligand binding cyt
28	61	16.8	603	21	Y32222	Granulocyte colony
29	61	16.8	771	12	R14256	Human GCSF recepto
30	61	16.8	783	12	R11741	Granulocyte colony
31	61	16.8	783	16	R77868	Human placental G-
32	61	16.8	783	18	W10485	Human granulocyte
33	61	16.8	800	18	W10486	Human granulocyte
34	61	16.8	801	12	R11742	Clone 25-1 encoded
35	61	16.8	836	12	R14255	Human GCSF recepto
36	61	16.8	863	12	R14257	Human GCSF recepto
37	60.5	16.6	131	18	W34327	Tomato mottle viru
38	60.5	16.6	447	21	Y87089	Human secreted pro
39	60.5	16.6	447	21	Y66744	Membrane-bound pro
40	60.5	16.6	447	21	Y50947	Human adult aorta
41	60.5	16.6	456	21	Y87190	Human secreted pro
42	59.5	16.3	404	18	W14436	OCH1 glycoprotein.
43	59.5	16.3	404	18	W14908	Pichia pastoris al
44	59.5	16.3	404	18	W01748	Alpha-1,6-mannosyl
45	58.5	16.1	619	13	R27651	Human calcium chan

ALIGNMENTS

RESULT 1
W34338
ID W34338 standard; Protein; 353 AA.
XX
AC W34338;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 protein.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93314.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

```
XX
PS   Example 5; Page 100-102; 132pp; English.
XX
CC   This sequence comprises the bean golden mosaic virus (BGMV) C1
CC   protein that is required for replication.  The invention involves
CC   production of transgenic plants containing DNA comprising C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking domain
CC   and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 223;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   2
W34332
ID   W34332 standard; Protein; 353 AA.
XX
AC   W34332;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93290.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 103-105; 132pp; English.
XX
CC   This protein comprises a control mutant of the bean golden mosaic
CC   virus (BGMV) C1 protein (see W34338) that is required for
CC   replication.  It is encoded by mutated C1 open reading frame
CC   BGAC190 (see T93290).  The invention involves production of
CC   transgenic plants containing DNA comprising geminivirus C1 or AC1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from BGMV, tomato mottle virus or tomato yellow leaf
CC   curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC   that have mutations in the highly conserved DNA-nicking and/or the
CC   NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 223;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   3
W34333
ID   W34333 standard; Protein; 353 AA.
XX
AC   W34333;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93291.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 107-109; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
```

```
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 223;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   4
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 223;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   5
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 223;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2.4e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169
```

```
Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             60.7%;  Score 221;  DB 8;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 4.3e-20;
  Matches   42;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:|||||||||||| |::||   :|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT   7
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             60.7%;  Score 221;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 4.4e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   8
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
```

```
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             60.7%;  Score 221;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 4.4e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
            | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   9
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             60.7%;  Score 221;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 4.4e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
            | || ||:|||||||| ||:||   :|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             60.7%;  Score 221;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 4.4e-20;
  Matches   40;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.7%;  Score 221;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 4.4e-20;
  Matches   40;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.7%;  Score 221;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 4.4e-20;
  Matches   40;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.7%;  Score 221;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 4.4e-20;
  Matches   40;  Conservative   15;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||:|||||||| |::||   :|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             58.4%;  Score 212.5;  DB 8;  Length 361;
  Best Local Similarity   64.6%;  Pred. No. 5.3e-19;
  Matches   42;  Conservative    9;  Mismatches   13;  Indels    1;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           | |||||||||||| |::||   ||||: || :|| |:||| |::|: |||||| || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             56.9%;  Score 207;  DB 19;  Length 362;
  Best Local Similarity   51.2%;  Pred. No. 2.6e-18;
  Matches   43;  Conservative   10;  Mismatches   15;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVDGRSARGGCQTSNDLLLEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | :| |||||||||  ||::||   ||:|: ||  || |:||| |: :: ||||||||||
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 RI----------------FDKTPE 69
           ||                ||: ||
Db     171 rifapplevfvcpfsssfdqvpe 194


Search completed: February  3, 2001, 02:15:30
Job time: 3437 sec
```

SEQ ID #9

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:29 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	356	98.3	352	1	VAL1_TGMV	P03567 tomato gold
2	256	70.7	361	1	VAL1_PYMVV	P27258 potato yell
3	243	67.1	358	1	VAL1_CLVK	P14982 cassava lat
4	243	67.1	358	1	VAL1_CLVN	P14972 cassava lat
5	236	65.2	362	1	VAL1_TYLCA	P36279 tomato yell
6	229	63.3	359	1	VAL1_TYLCU	P38609 tomato yell
7	225	62.2	349	1	VAL1_PHUV	Q06923 pepper huas
8	222	61.3	359	1	VAL1_TYLCM	P27260 tomato yell
9	220	60.8	353	1	VAL1_BGMV	P05175 bean golden
10	219	60.5	355	1	VAL1_ABMVW	P21947 abutilon mo
11	216	59.7	358	1	VAL1_BCTV	P14991 beet curly
12	216	59.7	361	1	VAL1_TMOV	Q06657 tomato mott
13	207	57.2	357	1	VAL1_TYLCV	P27259 tomato yell
14	125	34.5	347	1	VAL1_SLCV	P29048 squash leaf
15	64.5	17.8	299	1	Y175_HELPJ	Q9zmq7 helicobacte
16	62	17.1	295	1	VAL1_TYDVA	P31617 tobacco yel
17	61	16.9	1483	1	CYP1_YEAST	P12351 saccharomyc
18	60.5	16.7	1713	1	LMA3_HUMAN	Q16787 homo sapien
19	60	16.6	447	1	CDSA_DROME	P56079 d phosphati
20	59	16.3	1287	1	SKI2_YEAST	P35207 saccharomyc
21	58.5	16.2	297	1	RRPP_RABVP	P06747 rabies viru
22	58.5	16.2	299	1	Y175_HELPY	P56112 helicobacte
23	58.5	16.2	367	1	LHX4_MOUSE	P53776 mus musculu
24	58.5	16.2	589	1	YE67_HAEIN	P45221 haemophilus
25	58	16.0	207	1	IL6_MARMO	O35736 marmota mon
26	58	16.0	256	1	YKJ9_YEAST	P34247 saccharomyc
27	58	16.0	617	1	YACH_ECOLI	P36682 escherichia
28	58	16.0	874	1	SLAP_BACLI	P49052 bacillus li
29	57.5	15.9	136	1	Y452_CAEEL	O62250 caenorhabdi
30	57	15.7	355	1	CRTB_RHOSH	P54905 rhodobacter
31	56.5	15.6	428	1	GBA1_CANAL	P28868 candida alb
32	56.5	15.6	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
33	56	15.5	332	1	IPAD_SHIFL	P18013 shigella fl
34	56	15.5	336	1	Y625_METJA	Q58042 methanococc
35	56	15.5	513	1	HEM0_CHICK	P18080 gallus gall
36	56	15.5	795	1	LON_AQUAE	O66605 aquifex aeo
37	56	15.5	799	1	AFSK_STRCO	P54741 streptomyce
38	56	15.5	807	1	AFSK_STRGR	P54742 streptomyce
39	56	15.5	836	1	GCSR_HUMAN	Q99062 homo sapien
40	56	15.5	1610	1	CCAD_MESAU	Q99244 mesocricetu
41	56	15.5	2161	1	CCAD_HUMAN	Q01668 homo sapien
42	56	15.5	2203	1	CCAD_RAT	P27732 rattus norv
43	55.5	15.3	201	1	DRPF_CRAPL	P23283 craterostig
44	55.5	15.3	224	1	SP2R_BACSU	P39151 bacillus su
45	55.5	15.3	238	1	RIBB_ARCFU	O28173 archaeoglob

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 98.3%; Score 356; DB 1; Length 352;
Best Local Similarity 98.6%; Pred. No. 1.2e-32;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             70.7%;  Score 256;  DB 1;  Length 361;
  Best Local Similarity   69.6%;  Pred. No. 1.9e-21;
  Matches   48;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: || ||: ||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPEP 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             67.1%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   62.9%;  Pred. No. 5.5e-20;
  Matches   44;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||| |:|||: || ||| :||| :|:|:: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             67.1%;  Score 243;  DB 1;  Length 358;
  Best Local Similarity   62.9%;  Pred. No. 5.5e-20;
  Matches   44;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             65.2%;  Score 236;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 3.4e-19;
  Matches   47;  Conservative    9;  Mismatches   13;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||||: ||||||| |::||| |:||| || ||| ::||  |: |: |||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             63.3%;  Score 229;  DB 1;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 2e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             62.2%;  Score 225;  DB 1;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 5.5e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||: ||||||| |::|| |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             61.3%;  Score 222;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 1.2e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             60.8%;  Score 220;  DB 1;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 2e-17;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             60.5%;  Score 219;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||: ||||||| ||:||: |:|||| : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created) .
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.7%;  Score 216;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 5.7e-17;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||||||: ||||||| ||:||: |:|||| | ::|||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             59.7%;  Score 216;  DB 1;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.8e-17;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             57.2%;  Score 207;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 5.8e-16;
  Matches   40;  Conservative   10;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :|  |: ||||||| |::||| ||||| || ||| |::|| |: |: ||||||||||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             34.5%;  Score 125;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 8.5e-07;
  Matches   25;  Conservative   15;  Mismatches   22;  Indels    4;  Gaps    1;

Qy       5 GEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           |:::|:|    |     :|    |:|| |  ||| ||:   |: :: |:||| :|::|:|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
Y175_HELPJ
ID   Y175_HELPJ     STANDARD;      PRT;   299 AA.
AC   Q9ZMQ7;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN HP0175 PRECURSOR.
GN   HP0175 OR JHP0161.
OS   Helicobacter pylori J99 (Campylobacter pylori J99).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99120557; PubMed=9923682;
RA   Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA   Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA   Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA   Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA   Trust T.J.;
RT   "Genomic sequence comparison of two unrelated isolates of the human
RT   gastric pathogen Helicobacter pylori.";
RL   Nature 397:176-180(1999).
CC   -!- SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC       STRONG, TO C.JEJUNI CBF2.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001454; AAD05744.1; -.
DR   HSSP; Q13526; 1PIN.
DR   INTERPRO; IPR000297; -.
DR   PFAM; PF00639; Rotamase; 1.
DR   PROSITE; PS01096; PPIC_PPIASE_1; 1.
KW   Hypothetical protein; Isomerase; Rotamase; Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   CHAIN        22    299       HYPOTHETICAL PROTEIN HP0175.
FT   DOMAIN      154    253       PPIC-LIKE.
SQ   SEQUENCE   299 AA;  34040 MW;  9C037B1CD1110143 CRC64;


  Query Match             17.8%;  Score 64.5;  DB 1;  Length 299;
  Best Local Similarity   28.1%;  Pred. No. 4.2;
  Matches   18;  Conservative   14;  Mismatches   11;  Indels   21;  Gaps    2;
```

```
Qy      22 NDAAAEALNAS------------------SKEEALQIIREKIPEKYLFQFHNLNSNLDR 62
           |:| || || :                   :|::| :: : :|||| :  |:  |:| |:
Db      93 NEAKAEKLNQTPEFKAMMEAVKKQALVEFWAKKQAEEVKKIQIPEKEMQDFY--NANKDQ 150

Qy      63 IFDK 66
           :| |
Db     151 LFVK 154


Search completed: February  3, 2001, 02:24:29
Job time: 641 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:41 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	356	98.3	352	1	QQCVL1	AL1 protein - toma
2	256	70.7	361	1	QQCVPT	AL1 protein - pota
3	243	67.1	358	2	S07594	hypothetical prote
4	236	65.2	362	1	JQ1887	AL1 protein - toma
5	229	63.3	359	2	S39211	gene C1 protein -
6	225	62.2	349	2	JQ2300	replicase - pepper
7	225	62.2	349	2	S31875	AL1 protein - pepp
8	222	61.3	359	2	S22593	hypothetical prote
9	219	60.5	351	2	JQ2327	AL1 protein - Indi
10	219	60.5	355	1	QQCVW1	AV1 protein - abut
11	216	59.7	358	1	JQ1870	AL1 protein - toma
12	216	59.7	359	2	S39235	gene C1 protein -
13	216	59.7	385	2	S28360	AL1 protein - beet
14	209	57.7	360	2	S59885	replication-associ
15	207	57.2	357	1	QQCVC1	AL1 protein - toma
16	142	39.2	131	2	S45059	AC1 protein (clone
17	125	34.5	347	1	QQCVS1	AL1 protein - squa
18	68	18.8	587	2	JC1419	Fc gamma (IgG) rec
19	67	18.5	840	2	T36175	probable large ATP
20	64.5	17.8	299	2	B71967	probable peptidyl-
21	62.5	17.3	447	2	T12544	hypothetical prote
22	62	17.1	295	2	D42452	C1 protein - tobac
23	61.5	17.0	201	2	A81380	probable flagellar
24	61.5	17.0	481	2	A70091	probable phosphoes
25	61	16.9	1502	1	RGBYH1	CYC1/CYP3 transcri
26	60.5	16.7	1713	2	A55347	adhesive ligand ep
27	60.5	16.7	2137	2	T05244	hypothetical prote
28	60	16.6	160	2	GB2060	hypothetical prote
29	60	16.6	447	2	S52437	CDP-diacylglycerol
30	59	16.3	397	2	B71078	probable NADH oxid
31	59	16.3	416	2	A82892	hypothetical prote
32	59	16.3	862	2	H82182	conserved hypothet
33	59	16.3	1287	2	S55954	viral mRNA transla
34	59	16.3	1341	2	JG0166	LaMDR1 protein - L
35	58.5	16.2	230	2	S46332	homeotic protein G
36	58.5	16.2	297	1	MNVNRV	nonstructural prot
37	58.5	16.2	299	2	G64541	cell binding facto
38	58.5	16.2	589	2	D64125	hypothetical prote
39	58.5	16.2	1044	2	T43800	protein kinase bub
40	58.5	16.2	1265	2	T47626	structural mainten
41	58	16.0	245	2	S76632	hypothetical prote
42	58	16.0	256	2	S37926	hypothetical prote
43	58	16.0	617	2	E64734	yacH protein - Esc
44	58	16.0	874	2	JC4930	S-layer protein pr
45	57.5	15.9	136	2	T22240	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 98.3%; Score 356; DB 1; Length 352;
Best Local Similarity 98.6%; Pred. No. 1.9e-32;
Matches 69; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             70.7%;  Score 256;  DB 1;  Length 361;
  Best Local Similarity   69.6%;  Pred. No. 3.4e-21;
  Matches   48;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: || ||: ||||||| || |||||||||: :|| |::||:||:|||:|||:||| ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein

  Query Match             67.1%;  Score 243;  DB 2;  Length 358;
  Best Local Similarity   62.9%;  Pred. No. 9.8e-20;
  Matches   44;  Conservative   13;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||: ||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             65.2%;  Score 236;  DB 1;  Length 362;
  Best Local Similarity   55.3%;  Pred. No. 6.1e-19;
  Matches   47;  Conservative    9;  Mismatches   13;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || |||||: ||||||| |::||| |:||| || ||| ::|| |: |: |||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.3%;  Score 229;  DB 2;  Length 359;
  Best Local Similarity   60.9%;  Pred. No. 3.7e-18;
  Matches   42;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||: ||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179


RESULT   6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.2%;  Score 225;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1e-17;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||: ||||||| |::|| |:|||::| ||||||||:|: |:|:  |||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
```

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             62.2%;  Score 225;  DB 2;  Length 349;
  Best Local Similarity   60.0%;  Pred. No. 1e-17;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||||: ||||||| |::|| |:|||::| ||||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.3%;  Score 222;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 2.2e-17;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.5%;  Score 219;  DB 2;  Length 351;
  Best Local Similarity   62.7%;  Pred. No. 4.8e-17;
  Matches   42;  Conservative    9;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           || ||: ||||||| |::||| | |||: || |||:|:||  |  ||  ||:::|||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.5%;  Score 219;  DB 1;  Length 355;
  Best Local Similarity   58.6%;  Pred. No. 4.8e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||||: ||||||| ||:||: |:||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.7%;  Score 216;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 1.1e-16;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;
```

```
Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:||: ||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176
```

RESULT 12
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.7%;  Score 216;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.1e-16;
  Matches   39;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||: ||||||| ||:||| |:|:|| || ||| :|::  |  |:  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179
```

RESULT 13
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             59.7%;  Score 216;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.1e-16;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||||: ||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.7%;  Score 209;  DB 2;  Length 360;
  Best Local Similarity   60.6%;  Pred. No. 6.5e-16;
  Matches   40;  Conservative   12;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| ||: ||||||| |::||| ||||:|: ||  || |:|#| |: ::|||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.2%;  Score 207;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 1.1e-15;
  Matches   40;  Conservative   10;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      4 WGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|  |: ||||||| |::||| ||||: || ||| |::|| |: |:||||||:||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171
```

Search completed: February 3, 2001, 02:17:41
Job time: 2213 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:49 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	222	61.3	359	3	US-08-809-103B-2	Sequence 2, Appli
2	222	61.3	359	3	US-08-809-103B-4	Sequence 4, Appli
3	222	61.3	359	3	US-08-809-103B-6	Sequence 6, Appli
4	222	61.3	359	3	US-08-809-103B-8	Sequence 8, Appli
5	60.5	16.7	1713	3	US-08-600-982-24	Sequence 24, Appl
6	60.5	16.7	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
7	60	16.6	446	2	US-08-672-814D-11	Sequence 11, Appl
8	58	16.0	844	3	US-09-029-267-20	Sequence 20, Appl
9	56	15.5	454	3	US-08-446-100-31	Sequence 31, Appl
10	56	15.5	602	2	US-08-419-652-6	Sequence 6, Appli
11	56	15.5	771	1	US-07-923-976-6	Sequence 6, Appli
12	56	15.5	783	5	5422248-2	Patent No. 5422248
13	56	15.5	836	1	US-07-923-976-4	Sequence 4, Appli
14	56	15.5	863	1	US-07-923-976-8	Sequence 8, Appli
15	56	15.5	2161	1	US-07-745-206A-2	Sequence 2, Appli
16	56	15.5	2161	1	US-08-455-543A-49	Sequence 49, Appl
17	56	15.5	2161	1	US-08-455-543A-51	Sequence 51, Appl
18	56	15.5	2161	2	US-08-223-305C-49	Sequence 49, Appl
19	56	15.5	2161	2	US-08-223-305C-51	Sequence 51, Appl
20	56	15.5	2161	2	US-08-311-363-2	Sequence 2, Appli
21	54	14.9	454	3	US-08-446-100-30	Sequence 30, Appl
22	53.5	14.8	81	3	US-08-995-159-6	Sequence 6, Appli
23	53.5	14.8	312	1	US-08-414-625-2	Sequence 2, Appli
24	53	14.6	266	2	US-07-857-224B-40	Sequence 40, Appl
25	53	14.6	454	3	US-08-446-100-26	Sequence 26, Appl
26	53	14.6	454	3	US-08-446-100-27	Sequence 27, Appl
27	53	14.6	454	3	US-08-446-100-28	Sequence 28, Appl
28	53	14.6	454	3	US-08-446-100-29	Sequence 29, Appl
29	53	14.6	751	3	US-09-036-987A-24	Sequence 24, Appl
30	53	14.6	783	5	5231168-2	Patent No. 5231168
31	52.5	14.5	589	3	US-08-942-008-2	Sequence 2, Appli
32	52	14.4	439	1	US-07-762-681A-1	Sequence 1, Appli
33	51	14.1	203	2	US-08-531-525-23	Sequence 23, Appl
34	51	14.1	203	2	US-08-718-270A-23	Sequence 23, Appl
35	51	14.1	482	1	US-07-792-885A-1	Sequence 1, Appli
36	51	14.1	482	1	US-08-142-439A-7	Sequence 7, Appli
37	51	14.1	482	2	US-08-869-477-7	Sequence 7, Appli
38	51	14.1	824	1	US-08-221-750A-3	Sequence 3, Appli
39	50.5	14.0	103	1	US-08-552-142A-13	Sequence 13, Appl
40	50.5	14.0	374	2	US-08-928-692-51	Sequence 51, Appl
41	50.5	14.0	411	1	US-08-399-561-2	Sequence 2, Appli
42	50.5	14.0	3079	4	PCT-US94-00198-4	Sequence 4, Appli
43	50	13.8	165	4	PCT-US95-10075-4	Sequence 4, Appli
44	50	13.8	196	1	US-08-261-677-9	Sequence 9, Appli
45	50	13.8	196	1	US-08-384-556A-5	Sequence 5, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match             61.3%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;
```

```
Qy      2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             61.3%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             61.3%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             61.3%;  Score 222;  DB 3;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-21;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             16.7%;  Score 60.5;  DB 3;  Length 1713;
  Best Local Similarity   39.1%;  Pred. No. 40;
  Matches   18;  Conservative   10;  Mismatches   11;  Indels    7;  Gaps    3;

Qy      24 AAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFDK 66
             || :| |  ||: | ||| :|:| :| |       |:|| |:: ::
Db     568 AAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLNE 609


RESULT   6
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24


  Query Match             16.7%;  Score 60.5;  DB 4;  Length 1713;
  Best Local Similarity   39.1%;  Pred. No. 40;
```

```
  Matches   18;  Conservative   10;  Mismatches   11;  Indels    7;  Gaps    3;

Qy      24 AAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIFDK 66
           || :| |  ||: | ||| :|:| :| |       |:|| |:: ::
Db     568 AAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLLNE 609


RESULT   7
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
;  GENERAL INFORMATION:
;    APPLICANT:  Leung, David W.
;    APPLICANT:  Weeks, Reitha
;    TITLE OF INVENTION:  MAMMALIAN
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cell Therapeutics, Inc.
;      STREET:  201 Elliott Avenue West
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  U.S.A.
;      ZIP:  98119
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
;      COMPUTER:  PC Clone (486 microprocessor)
;      OPERATING SYSTEM:  MS-DOS Version 6.1, Windows NT
;      SOFTWARE:  WORD 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/672,814D
;      FILING DATE:  28-Jun-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Faciszewski, Stephen
;      REGISTRATION NUMBER:  36,131
;      REFERENCE/DOCKET NUMBER:  1803
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206)282-7100
;      TELEFAX:  (206)284-6206
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  446
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Drosophila
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    FEATURE:
;      NAME/KEY:  CDP-diacylglycerol synthase (CDS)
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
US-08-672-814D-11


  Query Match            16.6%;  Score 60;  DB 2;  Length 446;
  Best Local Similarity  34.0%;  Pred. No. 8.3;
  Matches   17;  Conservative    9;  Mismatches   22;  Indels    2;  Gaps    1;

Qy      20 TSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           : :||| :  :|:   : :    |||||:       |  || :  |||||
Db      20 SGSDAANKRNSAADSSDHVDSEEEKIPEEKFVD--ELAKNLPQGTDKTPE 67


RESULT   8
US-09-029-267-20
; Sequence 20, Application US/09029267
; Patent No. 6107057
;  GENERAL INFORMATION:
;    APPLICANT:  Crawford, Kenneth
;    APPLICANT:  Zaror, Isabel
;    APPLICANT:  Innis, Michael
;    TITLE OF INVENTION:  Pichia Secretory Leader for Protein
;    TITLE OF INVENTION:  Expression
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/029,267
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Chung, Ling-Fong
;      REGISTRATION NUMBER:  36,482
;      REFERENCE/DOCKET NUMBER:  1165.100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 601-2704
;      TELEFAX:  (510) 655-3542
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  844 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-029-267-20


  Query Match            16.0%;  Score 58;  DB 3;  Length 844;
  Best Local Similarity  27.5%;  Pred. No. 34;
  Matches   11;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      24 AAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           | :  |: :: |:||::  :  ||  : |  | :| :|::
Db     714 AHSTTLSVATYEQALEMSNQYAPEHLILQIENASSYVDQV 753


RESULT   9
US-08-446-100-31
; Sequence 31, Application US/08446100
; Patent No. 6001625
;  GENERAL INFORMATION:
;    APPLICANT:  Broomfield, Clarence A
;    APPLICANT:  Millard, Charles B
;    APPLICANT:  Lockridge, Oksana
;    TITLE OF INVENTION:  Site-Directed Mutagenesis of Esterases
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hendricks and Assoc.
;      STREET:  9669 A Main Street, P.O. Box 2509
;      CITY:  Fairfax
;      STATE:  VA
;      COUNTRY:  US
;      ZIP:  22031
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/446,100
;       FILING DATE:  19-MAY-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hendricks, Glenna
;       REGISTRATION NUMBER:  32,535
;       REFERENCE/DOCKET NUMBER:  broomfield
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 425-4250
;       TELEFAX:  (703) 425-2767
;   INFORMATION FOR SEQ ID NO:  31:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  454 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  protein
;     HYPOTHETICAL:  YES
;     ANTI-SENSE:  YES
;     FRAGMENT TYPE:  N-terminal
;     ORIGINAL SOURCE:
;       ORGANISM:  human esterases
US-08-446-100-31


  Query Match             15.5%;  Score 56;  DB 3;  Length 454;
  Best Local Similarity   32.3%;  Pred. No. 28;
  Matches   20;  Conservative   10;  Mismatches   22;  Indels   10;  Gaps    4;

Qy      17 GCQTSNDAA-AEALNASSKEEALQIIREKIPEKYLFQF-------HNLNSNLD-RIFDKT 67
           ||:|:  ||    |   ::|| |:    ||   ||: :       | | : :|  :  ||
Db     160 GCKTTTSAAMVHCLRQKTEEELLETTL-KIGNSYLWTYRETQRESHLLGTVIDGMLLLKT 218

Qy      68 PE 69
           ||
Db     219 PE 220


RESULT  10
US-08-419-652-6
; Sequence 6, Application US/08419652
; Patent No. 5831007
;  GENERAL INFORMATION:
;    APPLICANT:  Chua, Anne O
;    APPLICANT:  Gubler, Ulrich A
;    TITLE OF INVENTION:  INTERLEUKIN-12 RECEPTOR
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hoffmann-La Roche Inc.
;      STREET:  340 Kingsland Street
;      CITY:  Nutley
;      STATE:  New Jersey
;      COUNTRY:  United States of America
;      ZIP:  07110-1199
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/419,652
;      FILING DATE:  11-APR-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/248,532
;      FILING DATE:  31-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/094,713
;      FILING DATE:  19-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kass, Alan P
;      REGISTRATION NUMBER:  32142
;      REFERENCE/DOCKET NUMBER:  CD 9174
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (201) 235-4205
;      TELEFAX:  (201) 235-3500
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  602 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  1..602
;      OTHER INFORMATION:  /note= "Represents residues 98 to
;      OTHER INFORMATION:  731 of human granulocyte colony-stimulating
;      OTHER INFORMATION:  factor-receptor."
US-08-419-652-6


  Query Match             15.5%;  Score 56;  DB 2;  Length 602;
  Best Local Similarity   24.0%;  Pred. No. 40;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           | :    :|| |||  |:  : :    :           || |:|  : |:
Db      37 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 80


RESULT  11
US-07-923-976-6
; Sequence 6, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
```

```
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-6


  Query Match             15.5%;  Score 56;  DB 1;  Length 771;
  Best Local Similarity   24.0%;  Pred. No. 55;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy      7 FQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNL 56
          | :    :|| |||  |:  : :    :          || |:|  : |:
Db    156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT  12
5422248-2
;Patent No. 5422248
;    APPLICANT: SMITH, CRAIG A.;LARSEN, ALF D.;SIMS, JOHN E.;
;BENSON, CURTIS M.
;    TITLE OF INVENTION: DNA SEQUENCES ENCODING GRANULOCYTE-COLONY
;STIMULATING FACTOR RECEPTORS
;    NUMBER OF SEQUENCES: 6
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/6,183
;      FILING DATE: 15-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 587,329
;      FILING DATE: 24-SEP-1990
;      APPLICATION NUMBER: 522,952
;      FILING DATE: 03-APR-1990
;      APPLICATION NUMBER: 416,306
;      FILING DATE: 03-OCT-1989
;      APPLICATION NUMBER: 412,816
;      FILING DATE: 26-SEP-1989
;SEQ ID NO:2:
;      LENGTH: 783
5422248-2


  Query Match             15.5%;  Score 56;  DB 5;  Length 783;
  Best Local Similarity   24.0%;  Pred. No. 56;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy      7 FQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNL 56
          | :    :|| |||  |:  : :    :          || |:|  : |:
Db    156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT  13
US-07-923-976-4
; Sequence 4, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hellwege, James W.
;      REGISTRATION NUMBER:  28,808
;      REFERENCE/DOCKET NUMBER:  514853
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  703-415-1500
;      TELEFAX:  703-415-1508
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  836 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-923-976-4


  Query Match             15.5%;  Score 56;  DB 1;  Length 836;
  Best Local Similarity   24.0%;  Pred. No. 61;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy      7 FQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNL 56
          | :    :|| |||  |:  : :    :          || |:|  : |:
Db    156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT  14
US-07-923-976-8
; Sequence 8, Application US/07923976
; Patent No. 5574136
;  GENERAL INFORMATION:
;    APPLICANT:  Nagata, Shigekazu
;    APPLICANT:  Fukunaga, Rikiro
;    TITLE OF INVENTION:  DNA Encoding Granulocyte
;    TITLE OF INVENTION:  Colony-Stimulating Factor Receptor
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jones, Tullar & Cooper, P.C.
;      STREET:  P.O. Box 2266 Eads Station
;      CITY:  Arlington
;      STATE:  Virginia
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/923,976
;      FILING DATE:  19920922
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 74539/1990
;      FILING DATE:  23-MAR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 176629/1990
;      FILING DATE:  03-JUL-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP91/00375
;      FILING DATE:  22-MAR-1991
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME:  Hellwege, James W.
;       REGISTRATION NUMBER:  28,808
;       REFERENCE/DOCKET NUMBER:  514853
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-415-1500
;       TELEFAX:  703-415-1508
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  863 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-923-976-8


  Query Match             15.5%;  Score 56;  DB 1;  Length 863;
  Best Local Similarity   24.0%;  Pred. No. 64;
  Matches   12;  Conservative    9;  Mismatches   23;  Indels    6;  Gaps    1;

Qy       7 FQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNL 56
           |  :    :|| |||  |:  : :    :          || |:|  : |:
Db     156 FTLKSFKSRGNCQTQGDSILDCVPKDGQSHCC------IPRKHLLLYQNM 199


RESULT  15
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;     APPLICANT:  Harpold, Michael
;     APPLICANT:  Ellis, Steven
;     APPLICANT:  Williams, Mark
;     APPLICANT:  McCue, Ann
;     APPLICANT:  Feldman, Daniel
;     TITLE OF INVENTION:  Human Calcium Channel Compositions and
;     TITLE OF INVENTION:  Methods
;     NUMBER OF SEQUENCES:  32
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fitch, Even, Tabin & Flannery
;       STREET:  135 S. LaSalle
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  U.S.A.
;       ZIP:  60603
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/745,206A
;       FILING DATE:  19910815
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Feder, Scott B
;       REFERENCE/DOCKET NUMBER:  51504
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-372-7842
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2161 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             15.5%;  Score 56;  DB 1;  Length 2161;
  Best Local Similarity   51.9%;  Pred. No. 2.1e+02;
  Matches   14;  Conservative    4;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      22 NDAAAEALNASSKEEALQIIREKIPEK 48
           | | ||:|| : |||| :  |:||  |

Db     754 NLADAESLNTAQKEEAEEKERKKIARK 780


Search completed: February  3, 2001, 02:19:51
Job time: 1713 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:29 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-9
Perfect score: 362
Sequence: 1 TLVWGEFQVAGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	222	61.3	353	18	W34338	Bean golden mosaic
2	222	61.3	353	18	W34332	Bean golden mosaic
3	222	61.3	353	18	W34333	Bean golden mosaic
4	222	61.3	353	18	W34334	Bean golden mosaic
5	222	61.3	353	18	W34335	Bean golden mosaic
6	222	61.3	359	17	R88870	Sardinian tomato y
7	222	61.3	359	17	R88871	Sardinian tomato y
8	222	61.3	359	17	R88872	Sardinian tomato y
9	220	60.8	353	8	P70407	ORF 4 gene product
10	220	60.8	361	18	W34336	Tomato mottle viru
11	220	60.8	361	18	W34324	Tomato mottle viru
12	220	60.8	361	18	W34325	Tomato mottle viru
13	220	60.8	361	18	W34326	Tomato mottle viru
14	213.5	59.0	361	8	P70562	Product of ORF 4 f
15	208	57.5	362	19	W56495	Tobacco leaf curli
16	207	57.2	357	18	W34329	Tomato yellow leaf
17	207	57.2	357	18	W34330	Tomato yellow leaf
18	207	57.2	357	18	W34331	Tomato yellow leaf
19	199	55.0	357	18	W34337	Tomato yellow leaf
20	65	18.0	292	21	Y92317	Geminivirus Rep C1
21	65	18.0	335	21	Y92318	Mastrevirus Rep in
22	64.5	17.8	299	18	W55450	H. pylori ORF 02ae
23	64.5	17.8	512	19	W68473	HIV-1 strain YBF30
24	62.5	17.3	447	21	Y87089	Human secreted pro
25	62.5	17.3	447	21	Y66744	Membrane-bound pro
26	62.5	17.3	447	21	Y50947	Human adult aorta
27	62.5	17.3	456	21	Y87190	Human secreted pro
28	60.5	16.7	204	20	W89814	Protein encoded by
29	60.5	16.7	299	20	W89829	Protein encoded by
30	60.5	16.7	299	20	W89849	Antigen from clust
31	60.5	16.7	1713	16	R70148	Deduced sequence o
32	58.5	16.2	188	21	Y95043	Candida albicans p
33	58.5	16.2	297	8	P70318	Protein M1 encoded
34	58.5	16.2	299	19	W98322	H. pylori GHPO 136
35	58.5	16.2	299	19	W73034	Helicobacter pylor
36	58	16.0	844	18	W09773	Histidinol dehydro
37	57	15.7	108	16	R82847	Murine mutant cyto
38	56.5	15.6	249	20	W89822	Protein encoded by
39	56.5	15.6	269	20	W89860	Antigen 1 from clu
40	56.5	15.6	304	19	W98862	H. pylori GHPO 167
41	56	15.5	108	16	R82846	Murine mutant cyto
42	56	15.5	206	20	Y37577	Amino acid sequenc
43	56	15.5	213	17	W00648	Human G-CSF recept
44	56	15.5	333	17	R99141	Ligand binding cyt
45	56	15.5	454	21	Y49470	Human carboxyleste

ALIGNMENTS

RESULT 1
W34338
ID W34338 standard; Protein; 353 AA.
XX
AC W34338;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 protein.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93314.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

XX
PS Example 5; Page 100-102; 132pp; English.
XX
CC This sequence comprises the bean golden mosaic virus (BGMV) C1
CC protein that is required for replication. The invention involves
CC production of transgenic plants containing DNA comprising C1 or AC1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from BGMV, tomato mottle virus or tomato yellow leaf
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC that have mutations in the highly conserved DNA-nicking domain
CC and/or the NTP-binding domains.
XX
SQ Sequence 353 AA;

Query Match 61.3%; Score 222; DB 18; Length 353;
Best Local Similarity 60.0%; Pred. No. 7.2e-20;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179
```

RESULT 2
W34332
ID W34332 standard; Protein; 353 AA.
XX
AC W34332;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93290.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 103-105; 132pp; English.
XX
CC This protein comprises a control mutant of the bean golden mosaic
CC virus (BGMV) C1 protein (see W34338) that is required for
CC replication. It is encoded by mutated C1 open reading frame
CC BGAC190 (see T93290). The invention involves production of
CC transgenic plants containing DNA comprising geminivirus C1 or AC1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from BGMV, tomato mottle virus or tomato yellow leaf
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC that have mutations in the highly conserved DNA-nicking and/or the
CC NTP-binding domains.
XX
SQ Sequence 353 AA;

Query Match 61.3%; Score 222; DB 18; Length 353;
Best Local Similarity 60.0%; Pred. No. 7.2e-20;
Matches 42; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179
```

RESULT 3
W34333
ID W34333 standard; Protein; 353 AA.
XX
AC W34333;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93291.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 107-109; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminiviral
CC replication during infection. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGMV,
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the

```
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 222;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 7.2e-20;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   4
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 222;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 7.2e-20;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   5
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             61.3%;  Score 222;  DB 18;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 7.2e-20;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179
```

```
Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             61.3%;  Score 222;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-20;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   7
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             61.3%;  Score 222;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-20;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: ||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   8
```

```
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             61.3%;  Score 222;  DB 17;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 7.3e-20;
  Matches   40;  Conservative   11;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | || ||: |||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   9
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             60.8%;  Score 220;  DB 8;  Length 353;
  Best Local Similarity   60.0%;  Pred. No. 1.3e-19;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||| ||||||| |::||:.|:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             60.8%;  Score 220;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-19;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.8%;  Score 220;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-19;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.8%;  Score 220;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-19;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             60.8%;  Score 220;  DB 18;  Length 361;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-19;
  Matches   40;  Conservative   16;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:||: ||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  14
P70562
ID   P70562 standard; Protein; 361 AA.
XX
AC   P70562;
XX
DT   30-APR-1991  (first entry)
XX
DE   Product of ORF 4 from MYMV complementary strand (b).
XX
KW   Geminivirus.
XX
OS   Mungbean yellow mosaic virus.
XX
PN   JP62126982-A.
XX
PD   09-JUN-1987.
XX
PF   28-NOV-1985;   85JP-0266080.
XX
PR   28-NOV-1985;   85JP-0266080.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-196308/28.
DR   N-PSDB; N70895.
XX
PT   Novel DNA and hybrid DNA useful - as vector for recombinant work
PT   of plant gene.
XX
PS   Disclosure; Fig 8; 21pp; Japanese.
XX
CC   The sequence is encoded by ORF 4 which occurs on the complementary
CC   strand of the (b) molecule of the geminivirus.
CC   See also P70559-P70567.
```

```
XX
SQ   Sequence   361 AA;


  Query Match             59.0%;  Score 213.5;  DB 8;  Length 361;
  Best Local Similarity   66.2%;  Pred. No. 8.6e-19;
  Matches   43;  Conservative    9;  Mismatches   12;  Indels    1;  Gaps    1;

Qy       5 GEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
           | ||| ||||||| |::||| |||||: || :|| |:||| |::|: |||||| || |||
Db     114 gsfqvdgrsarggkqsandayaealnsgsklqalnilrekapkeyilqfhnlncnlsrif 173

Qy      65 -DKTP 68
            |: |
Db     174 sdevp 178


RESULT  15
W56495
ID   W56495 standard; Protein; 362 AA.
XX
AC   W56495;
XX
DT   11-AUG-1998  (first entry)
XX
DE   Tobacco leaf curling virus gene protein C1.
XX
KW   Tobacco leaf curling virus gene; TLCV; promoter; C1 protein.
XX
OS   Tobacco leaf curling virus.
XX
PN   JP10070982-A.
XX
PD   17-MAR-1998.
XX
PF   30-AUG-1996;   96JP-0230394.
XX
PR   30-AUG-1996;   96JP-0230394.
XX
PA   (NORQ ) NORINSUISANSHO KYUSHU NOGYO SHIKENJYO.
XX
DR   WPI; 1998-233630/21.
XX
PT   Tobacco leaf curling virus gene - useful for inserting into vectors
PT   for expression in, e.g. tomato plants
XX
PS   Claim 5; Fig 5; 9pp; Japanese.
XX
CC   This sequence represents the C1 protein encoded by the tobacco leaf
CC   curling virus (TLCV) gene of the invention. TLCV gene or its promoter can
CC   be inserted into a vector for expression in plants, e.g. tobacco and
CC   tomato. This sequence is believed to be encoded by the TLCV gene shown in
CC   V29761.
XX
SQ   Sequence   362 AA;


  Query Match             57.5%;  Score 208;  DB 19;  Length 362;
  Best Local Similarity   52.4%;  Pred. No. 4.2e-18;
  Matches   44;  Conservative   10;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       2 LVWGEFQVAGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | :| ||| ||||| ||::||| |||:|: ||  || |:||| |: :: ||||||||||
Db     111 ldfgvfqvdgrsarcccqsandayaeainsgskssalnilrekapkdfvlqfhnlnsnld 170

Qy      62 RI---------------FDKTPE 69
           ||               ||: ||
Db     171 rifapplevfvcpfssssfdqvpe 194


Search completed: February  3, 2001, 02:15:30
Job time: 3437 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:51 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	274	76.3	226	12	O09727	O09727 leonurus mo
2	274	76.3	226	12	Q9WHF6	Q9whf6 tomato mild
3	272	75.8	361	12	Q67574	Q67574 bean golden
4	270	75.2	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	270	75.2	364	12	Q9QS55	Q9qs55 sweet potat
6	267	74.4	185	12	Q98693	Q98693 sida golden
7	259	72.1	149	12	P88975	P88975 macroptiliu
8	259	72.1	233	12	Q9YLA4	Q9yla4 macroptiliu
9	258	71.9	234	12	O39180	O39180 potato yell
10	252	70.2	190	12	Q9Z089	Q9z089 tobacco lea
11	252	70.2	190	12	Q9Z084	Q9z084 tobacco lea
12	249	69.4	190	12	Q9W827	Q9w827 tobacco lea
13	249	69.4	208	12	Q9Z0C4	Q9z0c4 tobacco lea
14	249	69.4	353	12	O72692	O72692 beet curly
15	249	69.4	354	12	Q65438	Q65438 beet curly
16	247	68.8	208	12	Q9Z0C0	Q9z0c0 tobacco lea
17	247	68.8	208	12	Q9Z0B8	Q9z0b8 tobacco lea
18	246	68.5	203	12	Q9Z083	Q9z083 tobacco lea
19	245	68.2	208	12	Q9Z0B6	Q9z0b6 tobacco lea
20	245	68.2	358	12	Q65418	Q65418 beet curly
21	244	68.0	363	12	O73577	O73577 cotton leaf
22	238	66.3	190	12	Q9Z086	Q9z086 tobacco lea
23	238	66.3	362	12	O56816	O56816 chayote mos
24	237	66.0	208	12	Q9Z0A0	Q9z0a0 tobacco lea
25	237	66.0	362	12	Q9IN48	Q9in48 okra enatio
26	237	66.0	363	12	O72705	O72705 cotton leaf
27	237	66.0	363	12	O72719	O72719 cotton leaf
28	236	65.7	208	12	Q9Z0C6	Q9z0c6 tobacco lea
29	235	65.5	349	12	Q88888	Q88888 tomato pseu
30	234	65.2	359	12	Q9YZV4	Q9yzv4 tomato yell
31	234	65.2	359	12	Q9YZV2	Q9yzv2 tomato yell
32	234	65.2	359	12	Q9YUX7	Q9yux7 tomato yell
33	234	65.2	359	12	Q9YL27	Q9yl27 tomato yell
34	233	64.9	359	12	Q88942	Q88942 tomato yell
35	232	64.6	190	12	Q9Z0A7	Q9z0a7 tobacco lea
36	231	64.3	307	12	Q9IET7	Q9iet7 cotton leaf
37	231	64.3	361	12	O72723	O72723 cotton leaf
38	230	64.1	231	12	Q96620	Q96620 african tom
39	230	64.1	354	12	Q9IN42	Q9in42 south afric
40	229	63.8	153	12	O41475	O41475 wissadula g
41	229	63.8	358	12	Q9JEA2	Q9jea2 cassava gem
42	228	63.5	358	12	Q9WR17	Q9wr17 african cas
43	228	63.5	360	12	Q9YNW5	Q9ynw5 sida golden
44	226	63.0	351	12	Q9Q9R3	Q9q9r3 dicliptera
45	225	62.7	362	12	Q9YYT7	Q9yyt7 althea rose

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 76.3%; Score 274; DB 12; Length 226;
Best Local Similarity 77.1%; Pred. No. 2e-23;
Matches 54; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: |||||||||:||| || |||||||||  |  |||||:||:|||||||||||||| |
Db    111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy     61 AAIFDKTPEP 70
            || | |||
Db    171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             76.3%;  Score 274;  DB 12;  Length 226;
  Best Local Similarity   72.9%;  Pred. No. 2e-23;
  Matches   51;  Conservative   12;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:|||||||| ||:||||||||||||||||::||:||:|:::|||:|||:| |
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 AAIFDKTPEP 70
             || | |||
Db     171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             75.8%;  Score 272;  DB 12;  Length 361;
  Best Local Similarity   76.1%;  Pred. No. 5.7e-23;
  Matches   51;  Conservative    9;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           || |||||||||| ||:||||:||||||||||||:|||:||:|||:|||:|||:| |  |
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             75.2%;  Score 270;  DB 12;  Length 225;
  Best Local Similarity   76.1%;  Pred. No. 5.7e-23;
  Matches   51;  Conservative    8;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           || ||:|||||||| || ||||:||||||||||||:|||:||:|||:|||:|||:| |  |
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             75.2%;  Score 270;  DB 12;  Length 364;
  Best Local Similarity   79.4%;  Pred. No. 9.8e-23;
  Matches   54;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: |||||||||||||| ||:|||||||||| ||| |||||||:|||||||||| | |
Db    110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy     61 AAIFDKTP 68
            ||   |
Db    170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             74.4%;  Score 267;  DB 12;  Length 185;
  Best Local Similarity   70.0%;  Pred. No. 9.9e-23;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: || ||:|||||||| ||:|||||||||: :|||:|||:|||||:|||:||||:|||:| :
Db     89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy     61 AAIFDKTPEP 70
            || | |||
Db    149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             72.1%;  Score 259;  DB 12;  Length 149;
  Best Local Similarity   68.6%;  Pred. No. 6.3e-22;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: || ||:|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:| |
Db     52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy     61 AAIFDKTPEP 70
            || | |||
Db    112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             72.1%;  Score 259;  DB 12;  Length 233;
  Best Local Similarity   68.6%;  Pred. No. 1e-21;
```

```
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          ~ |: || ||:||||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:| |
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 AAIFDKTPEP 70
             || | |||
Db    170 DRIFMKDPEP 179


RESULT   9
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             71.9%;  Score 258;  DB 12;  Length 234;
  Best Local Similarity   68.6%;  Pred. No. 1.4e-21;
  Matches   48;  Conservative   12;  Mismatches   10;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: || ||:||||||||| ||:|||||||||: :||||::||:||:|||:|||:|||:| |
Db    110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy     61 AAIFDKTPEP 70
             || | | |
Db    170 DRIFMKAPNP 179


RESULT  10
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             70.2%;  Score 252;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 5.2e-21;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          || || ||:|||||||||| :||| ||||||||| ||| |||||:|:|:::||:|||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 ---------------AAIFDKTPE 69
                          |: ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             70.2%;  Score 252;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 5.2e-21;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          || || ||:|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 ---------------AAIFDKTPE 69
                          |: ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             69.4%;  Score 249;  DB 12;  Length 190;
  Best Local Similarity   58.8%;  Pred. No. 1.1e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || || ||||||||||||| :||| ||||||:|| ||| |||||:|: ::||:||||| |
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 AAI----------------FDKTPE 69
             |                  ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  13
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             69.4%;  Score 249;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 1.3e-20;
  Matches   50;  Conservative    8;  Mismatches   11;  Indels   16;  Gaps    1;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           || || ||||||||||||| :||| ||||||:|| ||| |||||:|: ::||:||||| |
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 AAI----------------FDKTPE 69
             |                  ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  14
O72692
ID   O72692      PRELIMINARY;      PRT;   353 AA.
AC   O72692;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   C1 PROTEIN.
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Briddon R.W.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X97203; CAA65841.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   353 AA;  40207 MW;  67A7ECFA3C2F4926 CRC64;


  Query Match             69.4%;  Score 249;  DB 12;  Length 353;
  Best Local Similarity   74.2%;  Pred. No. 2.3e-20;
  Matches   49;  Conservative    6;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           |||||||||||||| ||:|||||||||| :  |||||||||:||||:||:|||   | ||
Db     109 WGEFQVDGRSARGGQQTANDAAAEALNAGNALEALQIIREKLPEKYIFQYHNLKPNLEAI 168

Qy      64 FDKTPE 69
           |   |:
Db     169 FLPPPD 174


RESULT  15
Q65438
ID   Q65438      PRELIMINARY;      PRT;   354 AA.
AC   Q65438;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ORF L1.
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
OX   NCBI_TaxID=10840;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CFH;
RX   MEDLINE=94220740; PubMed=8167369;
RA   Stenger D.C.;
RT   "Complete nucleotide sequence of the hypervirulent CFH strain of beet
RT   curly top virus.";
RL   Mol. Plant Microbe Interact. 7:154-157(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CFH;
RA   Stenger D.C.;
RL   Submitted (OCT-1993) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U02311; AAA20513.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   354 AA;  40434 MW;  4D495DAA4BA1BDDE CRC64;
```

Query Match 69.4%; Score 249; DB 12; Length 354;
Best Local Similarity 74.2%; Pred. No. 2.3e-20;
Matches 49; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

```
Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           |||||||||||||| ||:|||||||||| :  |||||||||:||||:||:|||   | ||
Db     109 WGEFQVDGRSARGGQQTANDAAAEALNAGNALEALQIIREKLPEKYIFQYHNLKPNLEAI 168

Qy      64 FDKTPE 69
             |  |:
Db     169 FLPPPD 174
```

Search completed: February 3, 2001, 02:22:52
Job time: 1889 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:28 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	342	95.3	352	1	VAL1_TGMV	P03567 tomato gold
2	242	67.4	361	1	VAL1_PYMVV	P27258 potato yell
3	229	63.8	358	1	VAL1_CLVK	P14982 cassava lat
4	229	63.8	358	1	VAL1_CLVN	P14972 cassava lat
5	222	61.8	362	1	VAL1_TYLCA	P36279 tomato yell
6	216	60.2	349	1	VAL1_PHUV	Q06923 pepper huas
7	215	59.9	359	1	VAL1_TYLCU	P38609 tomato yell
8	212	59.1	358	1	VAL1_BCTV	P14991 beet curly
9	211	58.8	353	1	VAL1_BGMV	P05175 bean golden
10	211	58.8	359	1	VAL1_TYLCM	P27260 tomato yell
11	210	58.5	355	1	VAL1_ABMVW	P21947 abutilon mo
12	207	57.7	361	1	VAL1_TMOV	Q06657 tomato mott
13	193	53.8	357	1	VAL1_TYLCV	P27259 tomato yell
14	107	29.8	347	1	VAL1_SLCV	P29048 squash leaf
15	63.5	17.7	1610	1	CCAD_MESAU	Q99244 mesocricetu
16	63.5	17.7	2161	1	CCAD_HUMAN	Q01668 homo sapien
17	63.5	17.7	2203	1	CCAD_RAT	P27732 rattus norv
18	61	17.0	1483	1	CYP1_YEAST	P12351 saccharomyc
19	59.5	16.6	224	1	SP2R_BACSU	P39151 bacillus su
20	59.5	16.6	4466	1	DYHC_TRIGR	P23098 tripneustes
21	59	16.4	387	1	VE2_HPV41	P27552 human papil
22	59	16.4	387	1	Y4PF_RHISN	P55615 rhizobium s
23	58.5	16.3	136	1	Y452_CAEEL	O62250 caenorhabdi
24	58.5	16.3	299	1	Y175_HELPJ	Q9zmq7 helicobacte
25	57.5	16.0	367	1	LHX4_MOUSE	P53776 mus musculu
26	57.5	16.0	1852	1	CCAS_CYPCA	P22316 cyprinus ca
27	57	15.9	247	1	YCP4_YEAST	P25349 saccharomyc
28	57	15.9	256	1	YKJ9_YEAST	P34247 saccharomyc
29	56.5	15.7	2190	1	CCAD_CHICK	O73700 gallus gall
30	56	15.6	98	1	ATPF_METJA	Q57671 methanococc
31	56	15.6	295	1	VAL1_CSMV	P18921 chloris str
32	56	15.6	295	1	VAL1_TYDVA	P31617 tobacco yel
33	56	15.6	836	1	GCSR_HUMAN	Q99062 homo sapien
34	55.5	15.5	266	1	ETFB_MYCLE	O33095 mycobacteri
35	55.5	15.5	266	1	ETFB_MYCTU	O53276 mycobacteri
36	55.5	15.5	428	1	GBA1_CANAL	P28868 candida alb
37	55.5	15.5	887	1	MTP_BOVIN	P55156 bos taurus
38	55.5	15.5	1066	1	Z295_HUMAN	Q9ulj3 homo sapien
39	55.5	15.5	4466	1	DYHC_ANTCR	P39057 anthocidari
40	55	15.3	354	1	CARA_METJA	Q58425 methanococc
41	55	15.3	481	1	KNIR_DROVI	Q24753 drosophila
42	55	15.3	577	1	PT1_STRSL	P30299 streptococc
43	55	15.3	618	1	ORC2_DROME	Q24168 drosophila
44	55	15.3	898	1	UBPE_DROME	Q24574 drosophila
45	55	15.3	968	1	CC28_SCHPO	Q10752 schizosacch

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 95.3%; Score 342; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 1.3e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             67.4%;  Score 242;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 3e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: || ||:|||||||| || |||||||||: :|| |::||:||:|||:|||:|||:   |
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 AAIFDKTPE 69
             || | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             63.8%;  Score 229;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 8.9e-19;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: ||||||| |
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 AAIFDKTPEP 70
             || : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             63.8%;  Score 229;  DB 1;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 8.9e-19;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:||:|||||||| |::||| |:|||: || ||| :||| :|: :: ||||||| |
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy     61 AAIFDKTPEP 70
              || : | |
Db    169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             61.8%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 5.6e-18;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          || |||||:|||||||| |::||| |:||| || ||| ::|| |: |: ||||||| |
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 AAI----------------FDKTPE 69
             |                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.2%;  Score 216;  DB 1;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 2.6e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: |||||:|||||||| |::|| |:|||::| |||||||:|: |: :  ||||: |
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 AAIFDKTPEP 70
             ||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             59.9%;  Score 215;  DB 1;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 3.5e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 RVFQVPPAP 179


RESULT   8
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             59.1%;  Score 212;  DB 1;  Length 358;
  Best Local Similarity   55.7%;  Pred. No. 7.6e-17;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:|||||||| ||:||: |:|||||:| ::|||||::|: |: |  | ||| :
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 AAIFDKTPEP 70
             || : |:|
Db     170 QKIFQRPPDP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV      STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             58.8%;  Score 211;  DB 1;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 9.7e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: ||:|||||||||||| |::||: |:|||| | | || |::|: |:||: | ||: | |
Db     110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFFKVPEP 179


RESULT  10
VAL1_TYLCM
ID   VAL1_TYLCM      STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
```

```
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             58.8%;  Score 211;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 9.9e-17;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 KVFQVPPAP 179


RESULT  11
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             58.5%;  Score 210;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 1.3e-16;
  Matches   40;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             |  |||||:|||||||| ||:||: |:||||   : || |::|: |: |: | ||: | |
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFAKAPEP 179


RESULT  12
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             57.7%;  Score 207;  DB 1;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 2.8e-16;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
             |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: | |
Db     110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFAKAPEP 179


RESULT  13
VAL1_TYLCV
ID   VAL1_TYLCV     STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
```

```
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             53.8%;  Score 193;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 1.1e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
           :|  |:|||||||| |::||| |||||: || ||| |::|| |: |: ||||||:| |  |
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             29.8%;  Score 107;  DB 1;  Length 347;
  Best Local Similarity   34.8%;  Pred. No. 6.2e-05;
  Matches   23;  Conservative   13;  Mismatches   26;  Indels    4;  Gaps    1;

Qy       5 GEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAIF 64
           |:::| |    |     :|    |:|| |  ||| ||:   |: :: ;:||| : :  :|
Db     116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy      65 DKTPEP 70
             | |||
Db     172 QKPPEP 177


RESULT  15
CCAD_MESAU
ID   CCAD_MESAU     STANDARD;      PRT;  1610 AA.
AC   Q99244; Q99245;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   VOLTAGE-DEPENDENT L-TYPE CALCIUM CHANNEL ALPHA-1D SUBUNIT (CALCIUM
DE   CHANNEL, L TYPE, ALPHA-1 POLYPEPTIDE ISOFORM 2).
GN   CACNA1D OR CACNL1A2 OR CCHL1A2 OR CACH3 OR CACN4.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM HCA3A).
RC   TISSUE=INSULINOMA;
RX   MEDLINE=93149124; PubMed=1337146;
RA   Yaney G.C., Wheeler M.B., Wei X., Perez-Reyes E., Birnbaumer L.,
RA   Boyd A.E. III, Moss L.G.;
RT   "Cloning of a novel alpha 1-subunit of the voltage-dependent calcium
RT   channel from the beta-cell.";
RL   Mol. Endocrinol. 6:2143-2152(1992).
RN   [2]
RP   SEQUENCE OF 1146-1441 FROM N.A. (ISOFORMS CACH3B AND CACH3D).
RC   TISSUE=HEART;
RX   MEDLINE=91056091; PubMed=2173707;
RA   Perez-Reyes E., Wei X., Castellano A., Birnbaumer L.;
RT   "Molecular diversity of L-type calcium channels. Evidence for
RT   alternative splicing of the transcripts of three non-allelic genes.";
RL   J. Biol. Chem. 265:20430-20436(1990).
CC   -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC       ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC       IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC       CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC       CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
CC       GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
CC       CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC       GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
CC       PHENYLALKYLAMINES, BENZOTHIAZEPINES, AND BY OMEGA-AGATOXIN-IIIA
CC       (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
CC       GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC   -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC       COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC       IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC       FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC       SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC       CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC       LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
```

```
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; HCA3A (SHOWN HERE),
CC       CACH3B AND CACH3D; ARE PRODUCED BY ALTERNATIVE SPLICING. CACH3B
CC       IS IDENTICAL TO HCA3A IN THE FRAGMENT SEQUENCED.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, HEART AND SKELETAL MUSCLE.
CC   -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC       HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC       POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC       PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC       SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC   -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC       FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M57969; AAB59702.1; -.
DR   EMBL; M57970; AAA62807.1; -.
DR   INTERPRO; IPR000636; -.
DR   INTERPRO; IPR002077; -.
DR   INTERPRO; IPR003091; -.
DR   PFAM; PF00520; ion_trans; 4.
DR   PRINTS; PR00167; CACHANNEL.
DR   PRINTS; PR00169; KCHANNEL.
KW   Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW   Calcium channel; Glycoprotein; Repeat; Multigene family;
KW   Calcium-binding; Phosphorylation; Alternative splicing.
FT   REPEAT      112    408       I.
FT   REPEAT      508    754       II.
FT   REPEAT      871   1153       III.
FT   REPEAT     1190   1465       IV.
FT   DOMAIN        1    125       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    126    144       S1 OF REPEAT I (POTENTIAL).
FT   DOMAIN      145    162       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    163    182       S2 OF REPEAT I (POTENTIAL).
FT   DOMAIN      183    194       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    195    213       S3 OF REPEAT I (POTENTIAL).
FT   DOMAIN      214    234       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    235    253       S4 OF REPEAT I (POTENTIAL).
FT   DOMAIN      254    272       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    273    292       S5 OF REPEAT I (POTENTIAL).
FT   DOMAIN      293    380       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    381    405       S6 OF REPEAT I (POTENTIAL).
FT   DOMAIN      406    522       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    523    542       S1 OF REPEAT II (POTENTIAL).
FT   DOMAIN      543    557       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    558    576       S2 OF REPEAT II (POTENTIAL).
FT   DOMAIN      577    584       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    585    603       S3 OF REPEAT II (POTENTIAL).
FT   DOMAIN      604    613       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    614    632       S4 OF REPEAT II (POTENTIAL).
FT   DOMAIN      633    651       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    652    672       S5 OF REPEAT II (POTENTIAL).
FT   DOMAIN      673    726       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    727    751       S6 OF REPEAT II (POTENTIAL).
FT   DOMAIN      752    884       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    885    903       S1 OF REPEAT III (POTENTIAL).
FT   DOMAIN      904    919       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    920    939       S2 OF REPEAT III (POTENTIAL).
FT   DOMAIN      940    951       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    952    970       S3 OF REPEAT III (POTENTIAL).
FT   DOMAIN      971    976       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    977    996       S4 OF REPEAT III (POTENTIAL).
FT   DOMAIN      997   1015       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1016   1035       S5 OF REPEAT III (POTENTIAL).
FT   DOMAIN     1036   1125       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1126   1146       S6 OF REPEAT III (POTENTIAL).
FT   DOMAIN     1147   1203       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1204   1222       S1 OF REPEAT IV (POTENTIAL).
FT   DOMAIN     1223   1237       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1238   1257       S2 OF REPEAT IV (POTENTIAL).
FT   DOMAIN     1258   1264       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1265   1286       S3 OF REPEAT IV (POTENTIAL).
FT   DOMAIN     1287   1311       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1312   1331       S4 OF REPEAT IV (POTENTIAL).
FT   DOMAIN     1332   1350       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1351   1370       S5 OF REPEAT IV (POTENTIAL).
FT   DOMAIN     1371   1437       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   1438   1462       S6 OF REPEAT IV (POTENTIAL).
FT   DOMAIN     1463   1610       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN        1      6       POLY-MET.
FT   DOMAIN      652    658       POLY-LEU.
FT   DOMAIN      826    836       POLY-GLU.
FT   DOMAIN      428    445       BINDING TO THE BETA SUBUNIT (BY
FT                                SIMILARITY).
FT   SITE        363    363       CALCIUM ION SELECTIVITY AND PERMEABILITY
FT                                (BY SIMILARITY).
FT   SITE        704    704       CALCIUM ION SELECTIVITY AND PERMEABILITY
FT                                (BY SIMILARITY).
FT   SITE       1099   1099       CALCIUM ION SELECTIVITY AND PERMEABILITY
FT                                (BY SIMILARITY).
FT   SITE       1404   1404       CALCIUM ION SELECTIVITY AND PERMEABILITY
FT                                (BY SIMILARITY).
FT   BINDING    1073   1163       TO DIHYDROPYRIDINES (BY SIMILARITY).
FT   BINDING    1418   1484       TO DIHYDROPYRIDINES (BY SIMILARITY).
FT   BINDING    1430   1473       TO PHENYLALKYLAMINES (BY SIMILARITY).
FT   MOD_RES    1473   1473       PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT   CA_BIND    1491   1502       POTENTIAL.
FT   CARBOHYD    154    154       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    224    224       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    328    328       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC   1261   1303       GYFSDAWNTFDSLIVIGSIIDVALSEADPTESESLPLPTAT
FT                                PG -> HYFTDAWNTFDALIVVGSVVDIAITEVN (IN
FT                                ISOFORM CACH3D).
SQ   SEQUENCE   1610 AA;  182327 MW;  B3B2E3794D936F79 CRC64;
```

```
  Query Match             17.7%;  Score 63.5;  DB 1;  Length 1610;
  Best Local Similarity   28.4%;  Pred. No. 30;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     706 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 765

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     766 EEAEEKERKKIARK 779
```

Search completed: February 3, 2001, 02:24:28
Job time: 640 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:40 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	342	95.3	352	1	QQCVL1	AL1 protein - toma
2	242	67.4	361	1	QQCVPT	AL1 protein - pota
3	229	63.8	358	2	S07594	hypothetical prote
4	222	61.8	362	1	JQ1887	AL1 protein - toma
5	216	60.2	349	2	JQ2300	replicase - pepper
6	216	60.2	349	2	S31875	AL1 protein - pepp
7	215	59.9	359	2	S39211	gene C1 protein -
8	212	59.1	385	2	S28360	AL1 protein - beet
9	211	58.8	359	2	S22593	hypothetical prote
10	210	58.5	355	1	QQCVW1	AV1 protein - abut
11	207	57.7	358	1	JQ1870	AL1 protein - toma
12	205	57.1	351	2	JQ2327	AL1 protein - Indi
13	205	57.1	359	2	S39235	gene C1 protein -
14	195	54.3	360	2	S59885	replication-associ
15	193	53.8	357	1	QQCVC1	AL1 protein - toma
16	120	33.4	131	2	S45059	AC1 protein (clone
17	107	29.8	347	1	QQCVS1	AL1 protein - squa
18	70	19.5	587	2	JC1419	Fc gamma (IgG) rec
19	68	18.9	447	2	T12544	hypothetical prote
20	64	17.8	316	2	C82085	conserved hypothet
21	63.5	17.7	1229	2	T48959	kinesin-like prote
22	63.5	17.7	1610	2	A46227	voltage-dependent
23	63.5	17.7	1646	2	JH0422	voltage-dependent
24	63.5	17.7	2161	2	JH0564	calcium channel al
25	63.5	17.7	2181	2	A38198	calcium channel al
26	63.5	17.7	2203	2	T42742	voltage-dependent
27	62	17.3	245	2	S76632	hypothetical prote
28	61	17.0	1502	1	RGBYH1	CYC1/CYP3 transcri
29	60.5	16.9	840	2	T36175	probable large ATP
30	59.5	16.6	224	2	I40474	pro-sigma-E proces
31	59.5	16.6	4466	1	S17653	dynein beta heavy
32	59	16.4	370	2	T34050	hypothetical prote
33	59	16.4	384	1	W2WL41	E2 protein - human
34	58.5	16.3	136	2	T22240	hypothetical prote
35	58.5	16.3	299	2	B71967	probable peptidyl-
36	58.5	16.3	354	2	A75087	acetyl ornithine d
37	58	16.2	390	2	T30020	hypothetical prote
38	57.5	16.0	230	2	S46332	homeotic protein G
39	57.5	16.0	371	2	B69502	alcohol dehydrogen
40	57.5	16.0	419	2	T08683	hypothetical prote
41	57.5	16.0	1852	2	A37860	calcium channel pr
42	57.5	16.0	2137	2	T05244	hypothetical prote
43	57.5	16.0	5037	1	A54161	ryanodine-binding
44	57	15.9	247	2	S26733	hypothetical prote
45	57	15.9	256	2	S37926	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 95.3%; Score 342; DB 1; Length 352;
Best Local Similarity 95.7%; Pred. No. 2.1e-31;
Matches 67; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 AAIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             67.4%;  Score 242;  DB 1;  Length 361;
  Best Local Similarity   66.7%;  Pred. No. 5.5e-20;
  Matches   46;  Conservative   11;  Mismatches   12;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: || ||:||||||||| || ||||||||||: :|| |::||:||:|||:|||:|||:  |
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 AAIFDKTPE 69
           || | ||
Db    170 DRIFMKAPE 178


RESULT   3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             63.8%;  Score 229;  DB 2;  Length 358;
  Best Local Similarity   60.0%;  Pred. No. 1.6e-18;
  Matches   42;  Conservative   13;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||:|||||||| |::||| |:||: || ||| :||| :|: :: |||||| |
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 AAIFDKTPEP 70
           || : | |
Db    169 DRIFQEPPAP 178


RESULT   4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             61.8%;  Score 222;  DB 1;  Length 362;
  Best Local Similarity   52.9%;  Pred. No. 1e-17;
  Matches   45;  Conservative    9;  Mismatches   15;  Indels   16;  Gaps    1;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          || |||||:|||||||| |::||| |:|||  || ||| ::||  |: |: |||||||  |

Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYYVLQFHNLNSNL 169

Qy     61 AAI----------------FDKTPE 69
             |                  ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   5
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 4.8e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: |
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 AAIFDKTPEP 70
           ||   |||
Db    170 NRIFQTPPEP 179


RESULT   6
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and co
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.2%;  Score 216;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 4.8e-17;
  Matches   41;  Conservative   12;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||||:|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: |
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 AAIFDKTPEP 70
           ||   |||
Db    170 NRIFQTPPEP 179


RESULT   7
S39211
gene C1 protein - tomato yellow leaf curl virus
```

```
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from the
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.9%;  Score 215;  DB 2;  Length 359;
  Best Local Similarity   58.0%;  Pred. No. 6.5e-17;
  Matches   40;  Conservative    9;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 RVFQVPPAP 179


RESULT   8
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly to
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.1%;  Score 212;  DB 2;  Length 385;
  Best Local Similarity   55.7%;  Pred. No. 1.5e-16;
  Matches   39;  Conservative   15;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |: |||||:|||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :
Db     137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy      61 AAIFDKTPEP 70
             || : |:|
Db     197 QKIFQRPPDP 206


RESULT   9
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.8%;  Score 211;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.9e-16;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 KVFQVPPAP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.5%;  Score 210;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 2.4e-16;
  Matches   40;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
           |  |||||:|||||||| ||:||: |:||||   : || |::|: |: |: | ||: | |
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 AAIFDKTPEP 70
             || | |||
Db     170 ERIFAKAPEP 179


RESULT  11
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated f
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             57.7%;  Score 207;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 5.3e-16;
  Matches   38;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;
```

```
Qy      1 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSAL 60
          |: ||:||:|||||||| |::||: |:|||| | : || ::||: |: :: | ||: | |
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 AAIFDKTPEP 70
            || | |||
Db    167 ERIFAKAPEP 176
```

RESULT 12
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.1%;  Score 205;  DB 2;  Length 351;
  Best Local Similarity   59.7%;  Pred. No. 8.7e-16;
  Matches   40;  Conservative    9;  Mismatches   18;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
          || ||:|||||||| |::||| | |||: || |||:|:||  |  ||  ||:::| |  |
Db    113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy     64 FDKTPEP 70
          | | | |
Db    173 FTKPPPP 179
```

RESULT 13
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             57.1%;  Score 205;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 9e-16;
  Matches   38;  Conservative   11;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
          | || ||:|||||||| ||:||| |:|:|| || ||| :|::  |  |:  |||::| |
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 AIFDKTPEP 70
           :|   | |
Db    171 KVFQVPPAP 179
```

RESULT 14
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             54.3%;  Score 195;  DB 2;  Length 360;
  Best Local Similarity   57.6%;  Pred. No. 1.2e-14;
  Matches   38;  Conservative   12;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
          :| ||:|||||||| |::||| |||:|: ||  || |:||| |: :: |||||: |  |
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          |  : |
Db    173 FTPSAE 178
```

RESULT 15
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein

```
  Query Match             53.8%;  Score 193;  DB 1;  Length 357;
  Best Local Similarity   62.3%;  Pred. No. 2.1e-14;
  Matches   38;  Conservative   10;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      4 WGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALAAI 63
          :|  |:|||||||| |::||| |||||: || ||| |::|| |: |: |||||:| |  |
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171
```

Search completed: February 3, 2001, 02:17:41
Job time: 2213 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:47 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-7
Perfect score: 359
Sequence: 1 TLVWGEFQVDGRSARGGCQT.....FQFHNLNSALAAIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	211	58.8	359	3	US-08-809-103B-2	Sequence 2, Appli
2	211	58.8	359	3	US-08-809-103B-4	Sequence 4, Appli
3	211	58.8	359	3	US-08-809-103B-6	Sequence 6, Appli
4	211	58.8	359	3	US-08-809-103B-8	Sequence 8, Appli
5	63.5	17.7	2161	1	US-07-745-206A-2	Sequence 2, Appli
6	63.5	17.7	2161	1	US-08-455-543A-49	Sequence 49, Appl
7	63.5	17.7	2161	1	US-08-455-543A-51	Sequence 51, Appl
8	63.5	17.7	2161	2	US-08-223-305C-49	Sequence 49, Appl
9	63.5	17.7	2161	2	US-08-223-305C-51	Sequence 51, Appl
10	63.5	17.7	2161	2	US-08-311-363-2	Sequence 2, Appli
11	58	16.2	374	2	US-08-928-692-51	Sequence 51, Appl
12	57	15.9	454	3	US-08-446-100-26	Sequence 26, Appl
13	57	15.9	454	3	US-08-446-100-27	Sequence 27, Appl
14	57	15.9	454	3	US-08-446-100-28	Sequence 28, Appl
15	57	15.9	454	3	US-08-446-100-29	Sequence 29, Appl
16	57	15.9	454	3	US-08-446-100-30	Sequence 30, Appl
17	56	15.6	602	2	US-08-419-652-6	Sequence 6, Appli
18	56	15.6	771	1	US-07-923-976-6	Sequence 6, Appli
19	56	15.6	783	5	5422248-2	Patent No. 5422248
20	56	15.6	836	1	US-07-923-976-4	Sequence 4, Appli
21	56	15.6	863	1	US-07-923-976-8	Sequence 8, Appli
22	55.5	15.5	860	1	US-08-117-362-3	Sequence 3, Appli
23	55.5	15.5	860	1	US-08-486-924-3	Sequence 3, Appli
24	55	15.3	432	1	US-08-167-919A-11	Sequence 11, Appl
25	55	15.3	432	3	US-08-715-106-11	Sequence 11, Appl
26	55	15.3	454	3	US-08-446-100-31	Sequence 31, Appl
27	54.5	15.2	1713	3	US-08-600-982-24	Sequence 24, Appl
28	54.5	15.2	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
29	54	15.0	751	3	US-09-036-987A-24	Sequence 24, Appl
30	53.5	14.9	373	3	US-08-258-287B-42	Sequence 42, Appl
31	53.5	14.9	373	3	US-08-368-704C-42	Sequence 42, Appl
32	53.5	14.9	3898	2	US-08-876-991-2	Sequence 2, Appli
33	53.5	14.9	3898	2	US-09-059-853-2	Sequence 2, Appli
34	53	14.8	284	2	US-08-424-641B-5	Sequence 5, Appli
35	53	14.8	284	2	US-08-820-980-5	Sequence 5, Appli
36	53	14.8	284	2	US-08-826-439-5	Sequence 5, Appli
37	53	14.8	824	1	US-08-221-750A-3	Sequence 3, Appli
38	52.5	14.6	480	1	US-07-803-636A-2	Sequence 2, Appli
39	52	14.5	244	1	US-08-910-973-22	Sequence 22, Appl
40	52	14.5	439	1	US-07-762-681A-1	Sequence 1, Appli
41	52	14.5	482	1	US-07-792-885A-1	Sequence 1, Appli
42	52	14.5	482	1	US-08-142-439A-7	Sequence 7, Appli
43	52	14.5	482	2	US-08-869-477-7	Sequence 7, Appli
44	52	14.5	1012	2	US-08-680-326-34	Sequence 34, Appl
45	51	14.2	237	1	US-08-910-973-13	Sequence 13, Appl

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match            58.8%;  Score 211;  DB 3;  Length 359;
  Best Local Similarity  56.5%;  Pred. No. 2.6e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;
```

```
Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 AIFDKTPEP 70
           :|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             58.8%;  Score 211;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.6e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 AIFDKTPEP 70
           :|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             58.8%;  Score 211;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.6e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
          | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 AIFDKTPEP 70
           :|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             58.8%;  Score 211;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 2.6e-20;
  Matches   39;  Conservative   10;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 LVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSALA 61
           | || ||:|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|| |
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 AIFDKTPEP 70
             :|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             17.7%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   6
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
```

```
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             17.7%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |    ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   7
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match             17.7%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |    ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   8
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             17.7%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   9
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             17.7%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  10
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,363
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-51506
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             17.7%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 15;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEFQVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                             | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  11
US-08-928-692-51
; Sequence 51, Application US/08928692
; Patent No. 5958727
;  GENERAL INFORMATION:
;    APPLICANT:  Brody, Howard
;    APPLICANT:  Yaver, Deborah S.
;    APPLICANT:  Lamsa, Michael
;    APPLICANT:  Hansen, Kim
;    TITLE OF INVENTION:  Methods for Modifying the Production of
;    TITLE OF INVENTION:  a Polypeptide
;    NUMBER OF SEQUENCES:  80
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/928,692
;      FILING DATE:  12-SEPT-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  4944.200-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  374 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5958727e
US-08-928-692-51


  Query Match             16.2%;  Score 58;  DB 2;  Length 374;
  Best Local Similarity   35.3%;  Pred. No. 8.3;
  Matches   18;  Conservative    5;  Mismatches   12;  Indels   16;  Gaps    2;

Qy      22 NDAAAEALNASSKE---EALQIIREKIPEKYLFQFHNLNSALAAIFDKTPE 69
           :||  ||| ||:|   :| ||:| | :             | |  :|  ||
Db     181 SDAQREALTASAKAAGLEVLQLIHEPV----------AAALAYDARPE 218


RESULT  12
US-08-446-100-26
; Sequence 26, Application US/08446100
; Patent No. 6001625
;  GENERAL INFORMATION:
;    APPLICANT:  Broomfield, Clarence A
;    APPLICANT:  Millard, Charles B
;    APPLICANT:  Lockridge, Oksana
;    TITLE OF INVENTION:  Site-Directed Mutagenesis of Esterases
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hendricks and Assoc.
;      STREET:  9669 A Main Street, P.O. Box 2509
;      CITY:  Fairfax
;      STATE:  VA
;      COUNTRY:  US
;      ZIP:  22031
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,100
;      FILING DATE:  19-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hendricks, Glenna
;      REGISTRATION NUMBER:  32,535
;      REFERENCE/DOCKET NUMBER:  broomfield
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 425-4250
;      TELEFAX:  (703) 425-2767
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  454 amino acids
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  human esterases
US-08-446-100-26


  Query Match             15.9%;  Score 57;  DB 3;  Length 454;
  Best Local Similarity   30.6%;  Pred. No. 15;
  Matches   19;  Conservative   10;  Mismatches   23;  Indels   10;  Gaps    4;

Qy      17 GCQTSNDAA-AEALNASSKEEALQIIREKIPEKYLFQFHNL---NSALAAIFD-----KT 67
           ||:|:  ||    |   ::|| |:    ||   ||: :      :: |  : |     ||
Db     160 GCKTTTSAAMVHCLRQKTEEELLETTL-KIGNSYLWTYRETQRESTLLGTVIDGMLLLKT 218

Qy      68 PE 69
           ||
Db     219 PE 220


RESULT  13
US-08-446-100-27
; Sequence 27, Application US/08446100
; Patent No. 6001625
;  GENERAL INFORMATION:
;    APPLICANT:  Broomfield, Clarence A
;    APPLICANT:  Millard, Charles B
;    APPLICANT:  Lockridge, Oksana
;    TITLE OF INVENTION:  Site-Directed Mutagenesis of Esterases
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hendricks and Assoc.
;      STREET:  9669 A Main Street, P.O. Box 2509
;      CITY:  Fairfax
;      STATE:  VA
;      COUNTRY:  US
;      ZIP:  22031
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,100
;      FILING DATE:  19-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hendricks, Glenna
;      REGISTRATION NUMBER:  32,535
;      REFERENCE/DOCKET NUMBER:  broomfield
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 425-4250
;      TELEFAX:  (703) 425-2767
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  454 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  human esterases
US-08-446-100-27
```

```
  Query Match             15.9%;  Score 57;  DB 3;  Length 454;
  Best Local Similarity   30.6%;  Pred. No. 15;
  Matches   19;  Conservative   10;  Mismatches   23;  Indels   10;  Gaps    4;

Qy      17 GCQTSNDAA-AEALNASSKEEALQIIREKIPEKYLFQFHNL---NSALAAIFD-----KT 67
           ||:|:  ||    |   ::|| |:    ||   ||: :      :: |  : |     ||
Db     160 GCKTTTSAAMVHCLRQKTEEELLETTL-KIGNSYLWTYRETQRESTLLGTVIDGMLLLKT 218

Qy      68 PE 69
           ||
Db     219 PE 220


RESULT  14
US-08-446-100-28
; Sequence 28, Application US/08446100
; Patent No. 6001625
;  GENERAL INFORMATION:
;    APPLICANT:  Broomfield, Clarence A
;    APPLICANT:  Millard, Charles B
;    APPLICANT:  Lockridge, Oksana
;    TITLE OF INVENTION:  Site-Directed Mutagenesis of Esterases
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hendricks and Assoc.
;      STREET:  9669 A Main Street, P.O. Box 2509
;      CITY:  Fairfax
;      STATE:  VA
;      COUNTRY:  US
;      ZIP:  22031
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,100
;      FILING DATE:  19-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hendricks, Glenna
;      REGISTRATION NUMBER:  32,535
;      REFERENCE/DOCKET NUMBER:  broomfield
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 425-4250
;      TELEFAX:  (703) 425-2767
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  454 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  human esterases
US-08-446-100-28


  Query Match             15.9%;  Score 57;  DB 3;  Length 454;
  Best Local Similarity   30.6%;  Pred. No. 15;
  Matches   19;  Conservative   10;  Mismatches   23;  Indels   10;  Gaps    4;

Qy      17 GCQTSNDAA-AEALNASSKEEALQIIREKIPEKYLFQFHNL---NSALAAIFD-----KT 67
           ||:|:  ||    |   ::|| |:    ||   ||: :      :: |  : |     ||
Db     160 GCKTTTSAAMVHCLRQKTEEELLETTL-KIGNSYLWTYRETQRESTLLGTVIDGMLLLKT 218

Qy      68 PE 69
           ||
Db     219 PE 220
```

```
RESULT 15
US-08-446-100-29
; Sequence 29, Application US/08446100
; Patent No. 6001625
;  GENERAL INFORMATION:
;    APPLICANT:  Broomfield, Clarence A
;    APPLICANT:  Millard, Charles B
;    APPLICANT:  Lockridge, Oksana
;    TITLE OF INVENTION:  Site-Directed Mutagenesis of Esterases
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hendricks and Assoc.
;      STREET:  9669 A Main Street, P.O. Box 2509
;      CITY:  Fairfax
;      STATE:  VA
;      COUNTRY:  US
;      ZIP:  22031
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/446,100
;      FILING DATE:  19-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hendricks, Glenna
;      REGISTRATION NUMBER:  32,535
;      REFERENCE/DOCKET NUMBER:  broomfield
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 425-4250
;      TELEFAX:  (703) 425-2767
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  454 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  YES
;    FRAGMENT TYPE:  N-terminal
;    ORIGINAL SOURCE:
;      ORGANISM:  human esterases
US-08-446-100-29


  Query Match             15.9%;  Score 57;  DB 3;  Length 454;
  Best Local Similarity   30.6%;  Pred. No. 15;
  Matches   19;  Conservative   10;  Mismatches   23;  Indels   10;  Gaps    4;

Qy      17 GCQTSNDAA-AEALNASSKEEALQIIREKIPEKYLFQFHNL---NSALAAIFD-----KT 67
           ||:|:  ||    |   ::|| |:    ||   ||: :      :: |  : |     ||
Db     160 GCKTTTSAAMVHCLRQKTEEELLETTL-KIGNSYLWTYRETQRESTLLGTVIDGMLLLKT 218

Qy      68 PE 69
           ||
Db     219 PE 220


Search completed: February  3, 2001, 02:19:48
Job time: 1710 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:22:52 ; Search time 180.33 Seconds
(without alignments)
45.497 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	282	78.1	226	12	O09727	O09727 leonurus mo
2	282	78.1	226	12	Q9WHF6	Q9whf6 tomato mild
3	280	77.6	361	12	Q67574	Q67574 bean golden
4	278	77.0	225	12	Q9QDB1	Q9qdb1 cowpea gold
5	278	77.0	364	12	Q9QS55	Q9qs55 sweet potat
6	275	76.2	185	12	Q98693	Q98693 sida golden
7	267	74.0	149	12	P88975	P88975 macroptiliu
8	267	74.0	233	12	Q9YLA4	Q9yla4 macroptiliu
9	261	72.3	234	12	O39180	O39180 potato yell
10	258	71.5	190	12	Q9Z089	Q9z089 tobacco lea
11	258	71.5	190	12	Q9Z084	Q9z084 tobacco lea
12	257	71.2	190	12	Q9W827	Q9w827 tobacco lea
13	257	71.2	208	12	Q9Z0C4	Q9z0c4 tobacco lea
14	255	70.6	208	12	Q9Z0C0	Q9z0c0 tobacco lea
15	255	70.6	208	12	Q9Z0B8	Q9z0b8 tobacco lea
16	253	70.1	208	12	Q9Z0B6	Q9z0b6 tobacco lea
17	252	69.8	203	12	Q9Z083	Q9z083 tobacco lea
18	252	69.8	363	12	O73577	O73577 cotton leaf
19	245	67.9	208	12	Q9Z0A0	Q9z0a0 tobacco lea
20	245	67.9	363	12	O72705	O72705 cotton leaf
21	245	67.9	363	12	O72719	O72719 cotton leaf
22	244	67.6	190	12	Q9Z086	Q9z086 tobacco lea
23	244	67.6	208	12	Q9Z0C6	Q9z0c6 tobacco lea
24	243	67.3	349	12	Q88888	Q88888 tomato pseu
25	243	67.3	362	12	O56816	O56816 chayote mos
26	242	67.0	359	12	Q9YZV4	Q9yzv4 tomato yell
27	242	67.0	359	12	Q9YZV2	Q9yzv2 tomato yell
28	242	67.0	359	12	Q9YUX7	Q9yux7 tomato yell
29	242	67.0	359	12	Q9YL27	Q9yl27 tomato yell
30	241	66.8	353	12	O72692	O72692 beet curly
31	241	66.8	354	12	Q65438	Q65438 beet curly
32	241	66.8	359	12	Q88942	Q88942 tomato yell
33	240	66.5	190	12	Q9Z0A7	Q9z0a7 tobacco lea
34	240	66.5	362	12	Q9IN48	Q9in48 okra enatio
35	239	66.2	307	12	Q9IET7	Q9iet7 cotton leaf
36	239	66.2	361	12	O72723	O72723 cotton leaf
37	238	65.9	231	12	Q96620	Q96620 african tom
38	238	65.9	354	12	Q9IN42	Q9in42 south afric
39	237	65.7	358	12	Q65418	Q65418 beet curly
40	237	65.7	358	12	Q9JEA2	Q9jea2 cassava gem
41	233	64.5	362	12	Q9YYT7	Q9yyt7 althea rose
42	233	64.5	363	12	O73494	O73494 okra yellow
43	232	64.3	351	12	Q9Q9R3	Q9q9r3 dicliptera
44	231	64.0	306	12	Q9IET1	Q9iet1 cotton leaf
45	231	64.0	358	12	Q9WR17	Q9wr17 african cas

ALIGNMENTS

RESULT 1
O09727
ID O09727 PRELIMINARY; PRT; 226 AA.
AC O09727;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE REP PROTEIN (FRAGMENT).
GN REP.
OS Leonurus mosaic virus.
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=58177;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEMV- BRAZIL 1;
RA Faria J.C., Maxwell D.P.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; U92532; AAB51157.1; -.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
FT NON_TER 226 226
SQ SEQUENCE 226 AA; 25617 MW; 73CDB6E766083FC5 CRC64;

Query Match 78.1%; Score 282; DB 12; Length 226;
Best Local Similarity 78.6%; Pred. No. 5.4e-23;
Matches 55; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  |||||:||| || ||||||||||  |  |||||:||:||||||||||||||||
Db     111 TVEWGEFQVDGRSSRGGQQTVNDAAAEALNAPDKRTALQIIKEKLPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180
```

RESULT 2
Q9WHF6

```
ID   Q9WHF6      PRELIMINARY;      PRT;   226 AA.
AC   Q9WHF6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   tomato mild mottle geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Unclassified Geminiviridae.
OX   NCBI_TaxID=92943;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HN96-H5KW;
RA   Nakhla M.K., Mejia L., Ramirez P., Karkashian J.P., Doyle M.M.,
RA   Maxwell D.P.;
RT   "Molecular characterization and DNA-based detection methods for
RT   vegetable-infecting geminiviruses in Central America.";
RL   Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF131071; AAD33471.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     226    226
SQ   SEQUENCE   226 AA;  25941 MW;  2EA4116712871A23 CRC64;


  Query Match             78.1%;  Score 282;  DB 12;  Length 226;
  Best Local Similarity   74.3%;  Pred. No. 5.4e-23;
  Matches   52;  Conservative   12;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| ||:||||||||||||||||::|||:||:|:::|||:|||:|||
Db     111 TIEWGEFQIDGRSARGGQQTANDAAAEALNASSKEEAMRIIKEKLPKEFLFQYHNLSSNL 170

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     171 DRIFAKAPEP 180


RESULT   3
Q67574
ID   Q67574      PRELIMINARY;      PRT;   361 AA.
AC   Q67574;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PUTATIVE REPLICATIVE PROTEIN.
GN   AL1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10839;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Hanson S.F., Morales F.J., Ahlquist P.G.,
RA   Maxwell D.P., Russell D.R.;
RL   Phytopathology 81:980-985(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Hidayat S.H., Martinez R.T., Leong S.A., Faria J.C.,
RA   Morales F.J., Maxwell D.P.;
RL   Plant Dis. 75:336-342(1991).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Gilbertson R.L., Faria J.C., Ahlquist P.G., Maxwell D.P.;
RL   Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; M88686; AAA46312.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
DR   PRODOM; PD000736; -; 1.
SQ   SEQUENCE   361 AA;  41041 MW;  0094C7ACAF06B788 CRC64;


  Query Match             77.6%;  Score 280;  DB 12;  Length 361;
  Best Local Similarity   77.6%;  Pred. No. 1.5e-22;
  Matches   52;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||   |||||||| ||:||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQVDGRSARGGQQTANDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |:|
Db     173 FTKAPDP 179


RESULT   4
Q9QDB1
ID   Q9QDB1      PRELIMINARY;      PRT;   225 AA.
AC   Q9QDB1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   cowpea golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=69263;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CGMV-BR;
RA   Faria J.C.;
RT   "Partial nucleotide sequence of cowpea golden mosaic geminivirus from
RT   Brazil.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF188708; AAF06318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     225    225
SQ   SEQUENCE   225 AA;  25766 MW;  1089CB6BD8D15B5D CRC64;


  Query Match             77.0%;  Score 278;  DB 12;  Length 225;
  Best Local Similarity   77.6%;  Pred. No. 1.5e-22;
  Matches   52;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||   :|||||||| || ||||:||||||||||||:|||:||:|||:|||:|||:||||||
Db     113 WGHFQIDGRSARGGQQTINDAASEALNASSKEEAMQIIKEKLPEKFLFQYHNLSSNLDRI 172

Qy      64 FDKTPEP 70
           | | |||
Db     173 FKKPPEP 179


RESULT   5
Q9QS55
ID   Q9QS55      PRELIMINARY;      PRT;   364 AA.
AC   Q9QS55;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION INITIATION PROTEIN AC1.
GN   AC1.
OS   sweet potato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=100755;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Lotrakul P., Valverde R.A., Clark C.A., Sim J., De La Torre R.;
RT   "Detection of a geminivirus infecting sweet potato in the United
RT   States.";
RL   Plant Dis. 82:1253-1257(1998).
DR   EMBL; AF104036; AAD47173.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
SQ   SEQUENCE   364 AA;  40680 MW;  5F79752431A09D6E CRC64;


  Query Match             77.0%;  Score 278;  DB 12;  Length 364;
  Best Local Similarity   80.9%;  Pred. No. 2.5e-22;
  Matches   55;  Conservative    3;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  ||||||||| ||:|||||||||| ||| ||||||||:|||||||||||| |||
Db     110 TITWGEFQVDGRSARGGQQTANDAAAEALNAGSKEAALQIIREKLPEKYLFQFHNLVSNL 169

Qy      61 DRIFDKTP 68
           ||||   |
Db     170 DRIFSPPP 177


RESULT   6
Q98693
ID   Q98693      PRELIMINARY;      PRT;   185 AA.
AC   Q98693;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REP PROTEIN (FRAGMENT).
GN   AC1.
OS   sida golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51034;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA;
RA   Roye M.E., McLaughlin W.A., Nakhla N.K., Maxwell D.P.;
RL   Plant Dis. 81:1251-1258(1997).
DR   EMBL; U67926; AAB97865.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     185    185
SQ   SEQUENCE   185 AA;  20975 MW;  3913850A025A5EE1 CRC64;


  Query Match             76.2%;  Score 275;  DB 12;  Length 185;
  Best Local Similarity   71.4%;  Pred. No. 2.5e-22;
  Matches   50;  Conservative   12;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||   :|||||||| ||:||||||||: :||:||:|||||:||:||||:|||:||:
Db      89 TIEWGVFQIDGRSARGGQQTANDAAAEALNSGTKEDALKIIREKLPERYLFQYHNLSSNI 148

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     149 DRIFSKPPEP 158


RESULT   7
P88975
ID   P88975      PRELIMINARY;      PRT;   149 AA.
AC   P88975;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
GN   AC1.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICAN;
RA   Roye M.E.;
RL   Thesis (1996), Biochemistry, University of the West Indies, Jamaica.
DR   EMBL; U75278; AAB36919.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     149    149
SQ   SEQUENCE   149 AA;  16785 MW;  E4CF5EED4C9CD508 CRC64;


  Query Match             74.0%;  Score 267;  DB 12;  Length 149;
  Best Local Similarity   70.0%;  Pred. No. 1.4e-21;
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||   :|||||||| ||||||||||||: :|| |::|:::||:|||:|||:|||:|||
Db      52 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 111

Qy      61 DRIFDKTPEP 70
           |||| | |||
Db     112 DRIFMKDPEP 121


RESULT   8
Q9YLA4
ID   Q9YLA4      PRELIMINARY;      PRT;   233 AA.
AC   Q9YLA4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   REPLICATION ASSOCIATED PROTEIN (FRAGMENT).
GN   REP.
OS   Macroptilium golden mosaic geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=51676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E.;
RT   "Genetic diversity and phylogeny of whitefly-transmitted geminiviruses
RT   from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAMAICA STRAIN 1;
RA   Roye M.E., McLaughlin W.A., Maxwell D.P.;
RT   "Molecular characterization of two distinct geminiviruses infecting M.
RT   lathyroides from Jamaica.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF098940; AAD17850.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     233    233
SQ   SEQUENCE   233 AA;  26355 MW;  AA490AF4D2166A02 CRC64;


  Query Match             74.0%;  Score 267;  DB 12;  Length 233;
  Best Local Similarity   70.0%;  Pred. No. 2.4e-21;
```

```
  Matches   49;  Conservative   12;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||   :|||||||| ||||||||||||: :|| |::|::||:|||:|||:|||:|||
Db    110 TIEWGVFQIDGRSARGGQQTSNDAAAEALNSGTKEAAMRIVKEKLPEKFLFQYHNLSSNL 169

Qy     61 DRIFDKTPEP 70
          |||| | |||
Db    170 DRIFMKDPEP 179


RESULT   9
O39180
ID   O39180      PRELIMINARY;      PRT;   234 AA.
AC   O39180;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   REPLICATION-ASSOCIATED PROTEIN (FRAGMENT).
OS   potato yellow mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10827;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TOMATO STRAIN;
RA   Guzman P., Arredondo C.R., Emmatty D., Portillo R.J., Gilbertson R.L.;
RL   Plant Dis. 81:312-312(1997).
DR   EMBL; AF026553; AAB82605.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER     234    234
SQ   SEQUENCE   234 AA;  26486 MW;  9ED8F0697105CD19 CRC64;


  Query Match             72.3%;  Score 261;  DB 12;  Length 234;
  Best Local Similarity   68.6%;  Pred. No. 1.1e-20;
  Matches   48;  Conservative   13;  Mismatches    9;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||   :|||||||| ||:|||||||||: :||||::||:|||:|||:|||:|:|
Db    110 TIEWGLFQIDGRSARGGQQTANDAAAEALNSGTKEEAMKIIKEKLPEKFLFQYHNLSSDL 169

Qy     61 DRIFDKTPEP 70
          |||| | | |
Db    170 DRIFMKAPNP 179


RESULT  10
Q9Z089
ID   Q9Z089      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z089;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA3-1,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001315; BAA34033.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21432 MW;  AAC093D1D1610FAD CRC64;


  Query Match             71.5%;  Score 258;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 1.8e-20;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || ||   :|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  11
Q9Z084
ID   Q9Z084      PRELIMINARY;      PRT;   190 AA.
AC   Q9Z084;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE YOKOHAMA5-2,
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YOKOHAMA5;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001318; BAA34039.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  AAC1C2943E3F01AD CRC64;


  Query Match             71.5%;  Score 258;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 1.8e-20;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || ||   :|||||||||| :||| ||||||||| ||| |||||:|: ::||:|||||||
Db     85 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKAEALAIIREKLPKDFIFQYHNLNSNL 144

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    145 DRIFAPPLEVFVCPFTASSFDQVPE 169


RESULT  12
Q9W827
ID   Q9W827      PRELIMINARY;      PRT;   190 AA.
AC   Q9W827;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 PROTEIN (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
```

```
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GORA1;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001303; BAA34010.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     190    190
SQ   SEQUENCE   190 AA;  21444 MW;  93C3742A8EBDB7EB CRC64;


  Query Match             71.2%;  Score 257;  DB 12;  Length 190;
  Best Local Similarity   60.0%;  Pred. No. 2.3e-20;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   |||||||||||| :||| ||||||:|| ||| |||||:|: ::||:|||||||
Db      85 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 144

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     145 DRIFAPPLEVFVCPFSNSSFDQVPE 169


RESULT  13
Q9Z0C4
ID   Q9Z0C4      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C4;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE ABURA3-1, PARTIAL
DE   AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ABURA3;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001294; BAA33992.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23526 MW;  249CC31D8729C72D CRC64;


  Query Match             71.2%;  Score 257;  DB 12;  Length 208;
  Best Local Similarity   60.0%;  Pred. No. 2.5e-20;
  Matches   51;  Conservative    8;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   |||||||||||| :||| ||||||:|| ||| |||||:|: ::|||:|||||||
Db     102 TLEWGTFQVDGRSARGGCQNANDACAEALNAASKAEALSIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSNSSFDQVPE 186


RESULT  14
Q9Z0C0
ID   Q9Z0C0      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0C0;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-1(B152),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB152;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001296; BAA33996.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23486 MW;  E301135F799C3DAD CRC64;


  Query Match             70.6%;  Score 255;  DB 12;  Length 208;
  Best Local Similarity   58.8%;  Pred. No. 4.1e-20;
  Matches   50;  Conservative    9;  Mismatches   10;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   :|||||||||| :||| |||||||| :|| |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186


RESULT  15
Q9Z0B8
ID   Q9Z0B8      PRELIMINARY;      PRT;   208 AA.
AC   Q9Z0B8;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C1 AND C4 GENES, CLONE AMG-2(B154),
DE   PARTIAL AND COMPLETE CDS (FRAGMENT).
GN   C1.
OS   tobacco leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=67762;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AMGB154;
RA   Ooi K., Ohshita S., Ishii I., Yahara T.;
RT   "Molecular phylogeny of geminivirus infecting wild plants in Japan.";
RL   J. Plant Res. 110:247-257(1997).
DR   EMBL; AB001297; BAA33998.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
FT   NON_TER       1      1
FT   NON_TER     208    208
SQ   SEQUENCE   208 AA;  23472 MW;  6290D0EF7C9956AA CRC64;
```

Query Match 70.6%; Score 255; DB 12; Length 208;
Best Local Similarity 58.8%; Pred. No. 4.1e-20;
Matches 50; Conservative 9; Mismatches 10; Indels 16; Gaps 1;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || ||   :|||||||||| :||| ||||||||| :|| |||||:|: ::||:|||||||
Db     102 TLEWGTFQIDGRSARGGCQNANDACAEALNASSKADALAIIREKLPKDFIFQYHNLNSNL 161

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     162 DRIFAPPLEVFVCPFSSSSFDQVPE 186
```

Search completed: February 3, 2001, 02:22:52
Job time: 1889 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:24:28 ; Search time 83.07 Seconds
(without alignments)
27.213 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	350	97.0	352	1	VAL1_TGMV	P03567 tomato gold
2	250	69.3	361	1	VAL1_PYMVV	P27258 potato yell
3	237	65.7	358	1	VAL1_CLVK	P14982 cassava lat
4	237	65.7	358	1	VAL1_CLVN	P14972 cassava lat
5	230	63.7	362	1	VAL1_TYLCA	P36279 tomato yell
6	223	61.8	359	1	VAL1_TYLCU	P38609 tomato yell
7	219	60.7	349	1	VAL1_PHUV	Q06923 pepper huas
8	216	59.8	359	1	VAL1_TYLCM	P27260 tomato yell
9	214	59.3	353	1	VAL1_BGMV	P05175 bean golden
10	213	59.0	355	1	VAL1_ABMVW	P21947 abutilon mo
11	212	58.7	357	1	VAL1_TYLCV	P27259 tomato yell
12	210	58.2	358	1	VAL1_BCTV	P14991 beet curly
13	210	58.2	361	1	VAL1_TMOV	Q06657 tomato mott
14	117	32.4	347	1	VAL1_SLCV	P29048 squash leaf
15	68.5	19.0	1713	1	LMA3_HUMAN	Q16787 homo sapien
16	64.5	17.9	299	1	Y175_HELPJ	Q9zmq7 helicobacte
17	63.5	17.6	1610	1	CCAD_MESAU	Q99244 mesocricetu
18	63.5	17.6	2161	1	CCAD_HUMAN	Q01668 homo sapien
19	63.5	17.6	2203	1	CCAD_RAT	P27732 rattus norv
20	62	17.2	295	1	VAL1_TYDVA	P31617 tobacco yel
21	61	16.9	1483	1	CYP1_YEAST	P12351 saccharomyc
22	60	16.6	447	1	CDSA_DROME	P56079 d phosphati
23	59.5	16.5	136	1	Y452_CAEEL	O62250 caenorhabdi
24	59.5	16.5	299	1	Y175_HELPY	P56112 helicobacte
25	58.5	16.2	297	1	RRPP_RABVP	P06747 rabies viru
26	58.5	16.2	334	1	G3P_BACST	P00362 bacillus st
27	58.5	16.2	335	1	G3P_BACCO	P15115 bacillus co
28	58.5	16.2	428	1	GBA1_CANAL	P28868 candida alb
29	58.5	16.2	589	1	YE67_HAEIN	P45221 haemophilus
30	58.5	16.2	1070	1	P11B_HUMAN	P42338 homo sapien
31	58	16.1	617	1	YACH_ECOLI	P36682 escherichia
32	58	16.1	874	1	SLAP_BACLI	P49052 bacillus li
33	57	15.8	200	1	VIP_CHICK	P48143 gallus gall
34	57	15.8	200	1	VIP_MELGA	P45644 meleagris g
35	57	15.8	247	1	YCP4_YEAST	P25349 saccharomyc
36	57	15.8	355	1	CRTB_RHOSH	P54905 rhodobacter
37	57	15.8	387	1	Y4PF_RHISN	P55615 rhizobium s
38	57	15.8	4639	1	DYHC_DROME	P37276 drosophila
39	56.5	15.7	470	1	RHSA_RHIME	Q9z3r2 rhizobium m
40	56.5	15.7	2190	1	CCAD_CHICK	O73700 gallus gall
41	56	15.5	336	1	Y625_METJA	Q58042 methanococc
42	56	15.5	513	1	HEM0_CHICK	P18080 gallus gall
43	56	15.5	536	1	ACEA_EMENI	P28298 emericella
44	56	15.5	934	1	C1TC_HUMAN	P11586 h c-1-tetra
45	55.5	15.4	224	1	SP2R_BACSU	P39151 bacillus su

ALIGNMENTS

RESULT 1
VAL1_TGMV
ID VAL1_TGMV STANDARD; PRT; 352 AA.
AC P03567;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 PROTEIN.
GN AC1.
OS Tomato golden mosaic virus (TGMV).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN [1]
RP SEQUENCE FROM N.A.
RA Hamilton W.D.O., Stein V.E., Coutts R.H.A., Buck K.W.;
RT "Complete nucleotide sequence of the infectious cloned DNA components
RT of tomato golden mosaic virus: potential coding regions and regulatory
RT sequences.";
RL EMBO J. 3:2197-2205(1984).
CC -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC ---
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC ---
DR EMBL; K02029; -; NOT_ANNOTATED_CDS.
DR PIR; A04170; QQCVL1.
DR INTERPRO; IPR001191; -.
DR INTERPRO; IPR001301; -.
DR PFAM; PF00799; Gemini_AL1; 1.
DR PRINTS; PR00227; GEMCOATAL1.
DR PRINTS; PR00228; GEMCOATCLVL1.
KW ATP-binding.
FT NP_BIND 223 230 ATP (POTENTIAL).
SQ SEQUENCE 352 AA; 40332 MW; C33C938E9644B4A4 CRC64;

Query Match 97.0%; Score 350; DB 1; Length 352;
Best Local Similarity 97.1%; Pred. No. 9.1e-32;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
VAL1_PYMVV
ID VAL1_PYMVV STANDARD; PRT; 361 AA.

```
AC   P27258;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN.
OS   Potato yellow mosaic virus (isolate Venezuela).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91311403; PubMed=1856690;
RA   Coutts R.H.A., Coffin R.S., Roberts E.J.F., Hamilton W.D.O.;
RT   "The nucleotide sequence of the infectious cloned DNA components of
RT   potato yellow mosaic virus.";
RL   J. Gen. Virol. 72:1515-1520(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00940; BAA00782.1; -.
DR   PIR; JU0364; QQCVPT.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   361 AA;  40850 MW;  5627A33BF1264383 CRC64;


  Query Match             69.3%;  Score 250;  DB 1;  Length 361;
  Best Local Similarity   68.1%;  Pred. No. 1.3e-20;
  Matches   47;  Conservative   11;  Mismatches   11;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||   :||||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db     110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy      61 DRIFDKTPE 69
           |||| | ||
Db     170 DRIFMKAPE 178


RESULT   3
VAL1_CLVK
ID   VAL1_CLVK      STANDARD;      PRT;   358 AA.
AC   P14982;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain West Kenyan 844).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Gay M.R.;
RT   "Nucleotide sequence of cassava latent virus DNA.";
RL   Nature 301:260-262(1983).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J02057; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40346 MW;  ED173E753EE92D69 CRC64;


  Query Match             65.7%;  Score 237;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 3.5e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   4
VAL1_CLVN
ID   VAL1_CLVN      STANDARD;      PRT;   358 AA.
AC   P14972;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (40.4 KDA PROTEIN).
GN   AC1.
OS   Cassava latent virus (strain Nigerian).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90174930; PubMed=2308831;
RA   Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT   "Nucleotide sequence of the infectious cloned DNA components of
RT   African cassava mosaic virus (Nigerian strain).";
RL   Nucleic Acids Res. 18:197-198(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X17095; CAA34953.1; -.
DR   PIR; S07594; S07594.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40435 MW;  1DB16BB0CB2D5E2C CRC64;


  Query Match             65.7%;  Score 237;  DB 1;  Length 358;
  Best Local Similarity   61.4%;  Pred. No. 3.5e-19;
  Matches   43;  Conservative   13;  Mismatches   14;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :|||||||| |::||| |:|||: || ||| :||| :|: :: |||||||||
Db     109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168
```

```
Qy      61 DRIFDKTPEP 70
           |||| : | |
Db     169 DRIFQEPPAP 178


RESULT   5
VAL1_TYLCA
ID   VAL1_TYLCA     STANDARD;      PRT;   362 AA.
AC   P36279;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Australia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93139778; PubMed=8423446;
RA   Dry I.B., Rigden J.E., Krake L.R., Mullineaux P.M., Rezaian M.A.;
RT   "Nucleotide sequence and genome organization of tomato leaf curl
RT   geminivirus.";
RL   J. Gen. Virol. 74:147-151(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; S53251; CAB30888.1; -.
DR   PIR; JQ1887; JQ1887.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   362 AA;  41197 MW;  343E7184B4704098 CRC64;


  Query Match             63.7%;  Score 230;  DB 1;  Length 362;
  Best Local Similarity   54.1%;  Pred. No. 2.1e-18;
  Matches   46;  Conservative    9;  Mismatches   14;  Indels   16;  Gaps    1;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           || |||  :|||||||| |::||| |:|||  || ||| ::||  |: |: ||||||||||
Db     110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy      61 DRI----------------FDKTPE 69
           |||                ||: ||
Db     170 DRIFTPPLEVYVSPFLSSSFDRVPE 194


RESULT   6
VAL1_TYLCU
ID   VAL1_TYLCU     STANDARD;      PRT;   359 AA.
AC   P38609;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Murcia) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94256836; PubMed=8198442;
RA   Noris E., Hidalgo E., Accotto G., Moriones E.;
RT   "High similarity among the tomato yellow leaf curl virus isolates
RT   from the west Mediterranean basin: the nucleotide sequence of an
RT   infectious clone from Spain.";
RL   Arch. Virol. 135:165-170(1994).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z25751; CAA81026.1; -.
DR   PIR; S39211; S39211.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  41065 MW;  2D170A51EF80A3EC CRC64;


  Query Match             61.8%;  Score 223;  DB 1;  Length 359;
  Best Local Similarity   59.4%;  Pred. No. 1.3e-17;
  Matches   41;  Conservative    9;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           |:|   | |
Db     171 RVFQVPPAP 179


RESULT   7
VAL1_PHUV
ID   VAL1_PHUV      STANDARD;      PRT;   349 AA.
AC   Q06923;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
OS   Pepper huasteco virus (PHV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94015007; PubMed=8409944;
RA   Torres-Pacheco I., Garzon-Tiznado J.A., Herrera-Estrella L.,
RA   Rivera-Bustamante R.F.;
RT   "Complete nucleotide sequence of pepper huasteco virus: analysis and
RT   comparison with bipartite geminiviruses.";
RL   J. Gen. Virol. 74:2225-2231(1993).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70418; CAA49856.1; -.
DR   PIR; S31875; S31875.
DR   PIR; JQ2300; JQ2300.
DR   INTERPRO; IPR001191; -.
```

```
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (BY SIMILARITY).
SQ   SEQUENCE   349 AA;  39722 MW;  D5F4E76CD56370F4 CRC64;


  Query Match             60.7%;  Score 219;  DB 1;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 3.5e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||  :|||||||| |::|| |:|||::| ||||||||:|: |: :  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179


RESULT   8
VAL1_TYLCM
ID   VAL1_TYLCM     STANDARD;      PRT;   359 AA.
AC   P27260;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (strain Marmande) (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92107660; PubMed=1840676;
RA   Kheyr-Pour A., Bendahmane M., Matzeit V., Accotto G.P., Crespi S.,
RA   Gronenborn B.;
RT   "Tomato yellow leaf curl virus from Sardinia is a
RT   whitefly-transmitted monopartite geminivirus.";
RL   Nucleic Acids Res. 19:6763-6769(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61153; CAA43466.1; -.
DR   PIR; S22593; S22593.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     220    227       ATP (POTENTIAL).
SQ   SEQUENCE   359 AA;  40733 MW;  9717B4A07C93EFA7 CRC64;


  Query Match             59.8%;  Score 216;  DB 1;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 7.7e-17;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   9
VAL1_BGMV
ID   VAL1_BGMV     STANDARD;      PRT;   353 AA.
AC   P05175;
DT   13-AUG-1987 (Rel. 05, Created)
DT   13-AUG-1987 (Rel. 05, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AL1 PROTEIN (40.2 KDA PROTEIN).
GN   AC1.
OS   Bean golden mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Howarth A.J., Caton J., Bossert M., Goodman R.M.;
RT   "Nucleotide sequence of bean golden mosaic virus and a model for gene
RT   regulation in geminiviruses.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:3572-3576(1985).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M10070; AAA46318.1; -.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   353 AA;  40190 MW;  80FA779DF6029A34 CRC64;


  Query Match             59.3%;  Score 214;  DB 1;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 1.3e-16;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 TIEWGQFQVDGRSARGGQQSANDSYAKALNADSIESALTILKEEQPKDYVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFFKVPEP 179


RESULT  10
VAL1_ABMVW
ID   VAL1_ABMVW     STANDARD;      PRT;   355 AA.
AC   P21947;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AC1.
OS   Abutilon mosaic virus (isolate West India).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91020984; PubMed=2219703;
RA   Frischmuth T., Zimmat G., Jeske H.;
RT   "The nucleotide sequence of abutilon mosaic virus reveals prokaryotic
RT   as well as eukaryotic features.";
RL   Virology 178:461-468(1990).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15983; -; NOT_ANNOTATED_CDS.
DR   PIR; A36214; QQCVW1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     221    228       ATP (POTENTIAL).
SQ   SEQUENCE   355 AA;  40257 MW;  16A2CA8A63251E95 CRC64;


  Query Match             59.0%;  Score 213;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 1.6e-16;
  Matches   40;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  |||  :|||||||| ||:||: |:|||||   : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
VAL1_TYLCV
ID   VAL1_TYLCV      STANDARD;      PRT;   357 AA.
AC   P27259;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN (C1 PROTEIN).
GN   C1.
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92024070; PubMed=1926771;
RA   Navot N., Pichersky E., Zeidan M., Zamir D., Czosnek H.;
RT   "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
RT   with a single genomic component.";
RL   Virology 185:151-161(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X15656; CAA33688.1; -.
DR   PIR; D40779; QQCVC1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     219    226       ATP (POTENTIAL).
SQ   SEQUENCE   357 AA;  40678 MW;  939AB68E1AB3B2A7 CRC64;


  Query Match             58.7%;  Score 212;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 2.1e-16;
  Matches   40;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           :| : :|||||||| |::||| ||||| |: ||| |::|| |: |: ||||:|||||
Db     111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy      64 F 64
           |
Db     171 F 171


RESULT  12
VAL1_BCTV
ID   VAL1_BCTV      STANDARD;      PRT;   358 AA.
AC   P14991;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   AL1 PROTEIN (40.8 KDA PROTEIN).
OS   Beet curly top virus (BCTV).
OC   Viruses; ssDNA viruses; Geminiviridae; Curtovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stanley J., Markham P.G., Callis R.J., Pinner M.S.;
RT   "The nucleotide sequence of an infectious clone of the geminivirus
RT   beet curly top virus.";
RL   EMBO J. 5:1761-1767(1986).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X04144; -; NOT_ANNOTATED_CDS.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (POTENTIAL).
SQ   SEQUENCE   358 AA;  40889 MW;  39A45FE3C0B9C333 CRC64;


  Query Match             58.2%;  Score 210;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 3.6e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :|
Db     110 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 169

Qy      61 DRIFDKTPEP 70
           :|| : |:|
Db     170 QKIFQRPPDP 179


RESULT  13
VAL1_TMOV
ID   VAL1_TMOV      STANDARD;      PRT;   361 AA.
AC   Q06657;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   AL1 PROTEIN.
GN   AL1.
```

```
OS   Tomato mottle virus (isolate Florida) (TMoV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93107858; PubMed=1469361;
RA   Abouzid A.M., Polston J.E., Hiebert E.;
RT   "The nucleotide sequence of tomato mottle virus, a new geminivirus
RT   isolated from tomatoes in Florida.";
RL   J. Gen. Virol. 73:3225-3229(1992).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L14460; AAC32414.1; -.
DR   PIR; JQ1870; JQ1870.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     222    229       ATP (BY SIMILARITY).
SQ   SEQUENCE   361 AA;  40516 MW;  8138B65CEEAC6950 CRC64;


  Query Match             58.2%;  Score 210;  DB 1;  Length 361;
  Best Local Similarity   54.3%;  Pred. No. 3.6e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  :||||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db    110 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 ERIFAKAPEP 179


RESULT  14
VAL1_SLCV
ID   VAL1_SLCV      STANDARD;      PRT;   347 AA.
AC   P29048;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   AL1 PROTEIN.
OS   Squash leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91082449; PubMed=1984668;
RA   Lazarowitz S.G., Lazdins I.B.;
RT   "Infectivity and complete nucleotide sequence of the cloned genomic
RT   components of a bipartite squash leaf curl geminivirus with a broad
RT   host range phenotype.";
RL   Virology 180:58-69(1991).
CC   -!- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M38183; AAC32410.1; ALT_INIT.
DR   PIR; C36785; QQCVS1.
DR   INTERPRO; IPR001191; -.
DR   INTERPRO; IPR001301; -.
DR   PFAM; PF00799; Gemini_AL1; 1.
DR   PRINTS; PR00227; GEMCOATAL1.
DR   PRINTS; PR00228; GEMCOATCLVL1.
KW   ATP-binding.
FT   NP_BIND     218    225       ATP (POTENTIAL).
SQ   SEQUENCE   347 AA;  39110 MW;  AFDAEBDDE122110E CRC64;


  Query Match             32.4%;  Score 117;  DB 1;  Length 347;
  Best Local Similarity   37.9%;  Pred. No. 7.8e-06;
  Matches   25;  Conservative   12;  Mismatches   25;  Indels    4;  Gaps    1;

Qy      5 GEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIF 64
          |:  | |    |     :|    |:|| |  ||| ||:   |: :: |  :||| :|:::|:|
Db    116 GQYKVSG----GSKSNKDDVYHNAVNAGSAGEALDIIKAGDPKTFIVNYHNLLANVERLF 171

Qy     65 DKTPEP 70
            | |||
Db    172 QKPPEP 177


RESULT  15
LMA3_HUMAN
ID   LMA3_HUMAN     STANDARD;      PRT;  1713 AA.
AC   Q16787; Q13679; Q13680;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   LAMININ ALPHA-3 CHAIN PRECURSOR (EPILIGRIN 170 KDA SUBUNIT) (E170).
GN   LAMA3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=KERATINOCYTES;
RX   MEDLINE=94357926; PubMed=8077230;
RA   Ryan M.C., Tizard R., Vandevanter D.R., Carter W.G.;
RT   "Cloning of the LamA3 gene encoding the alpha 3 chain of the adhesive
RT   ligand epiligrin. Expression in wound repair.";
RL   J. Biol. Chem. 269:22779-22787(1994).
RN   [2]
RP   SEQUENCE OF 1-1331 FROM N.A. (ISOFORMS A AND B).
RX   MEDLINE=96163880; PubMed=8586427;
RA   Vidal F., Baudoin C., Miquel C., Galliano M.-F., Christiano A.M.,
RA   Uitto J., Ortonne J.-P., Meneguzzi G.;
RT   "Cloning of the laminin alpha 3 chain gene (LAMA3) and identification
RT   of a homozygous deletion in a patient with Herlitz junctional
RT   epidermolysis bullosa.";
RL   Genomics 30:273-280(1995).
CC   -!- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC       IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC       CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC       WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC   -!- FUNCTION: LAMININ-5 IS THOUGHT TO BE INVOLVED IN (1) CELL ADHESION
CC       VIA INTEGRIN ALPHA-3/BETA-1 IN FOCAL ADHESION AND INTEGRIN ALPHA-
CC       6/BETA-4 IN HEMIDESMOSOMES, (2) SIGNAL TRANSDUCTION VIA TYROSINE
CC       PHOSPHORYLATION OF PP125-FAK AND P80, (3) DIFFERENTIATION OF
CC       KERATINOCYTES.
CC   -!- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC       DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC       TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC       COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC       THE ALPHA-3 CHAIN IS A SUBUNIT OF LAMININ-5 (EPILIGRIN/KALININ/
CC       NICEIN), AND POSSIBLY ALSO A COMPONENT OF LAMININ-6 (K-LAMININ)
CC       AND LAMININ-7 (KS-LAMININ).
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
CC       MEMBRANES (MAJOR COMPONENT).
CC   -!- ALTERNATIVE PRODUCTS: THE TWO ISOFORMS A AND B DIFFER IN THEIR N-
CC       TERMINUS. THE SEQUENCE SHOWN HERE IS THAT OF THE SMALLER VARIANT
```

```
CC       A.
CC   -!- TISSUE SPECIFICITY: SKIN; RESPIRATORY, URINARY, AND DIGESTIVE
CC       EPITHELIA AND IN OTHER SPECIALIZED TISSUES WITH PROMINENT
CC       SECRETORY OR PROTECTIVE FUNCTIONS. EPITHELIAL BASEMENT MEMBRANE,
CC       AND EPITHELIAL CELL TONGUE THAT MIGRATES INTO A WOUND BED. A
CC       DIFFERENTIAL AND FOCAL EXPRESSION OF THE ALPHA-3 CHAIN IS OBSERVED
CC       IN THE CNS.
CC   -!- INDUCTION: LAMININ-5 IS UP-REGULATED IN WOUND SITES OF HUMAN SKIN.
CC   -!- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC       WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC   -!- DOMAIN: DOMAIN G IS GLOBULAR.
CC   -!- DISEASE: JUNCTIONAL EPIDERMOLYSIS BULLOSA GRAVIS IS A BLISTERING
CC       DISORDER IN SKIN THAT IS CHARACTERIZED BY A SEPARATION OF BASAL
CC       CELLS FROM THE BASEMENT MEMBRANE DUE TO A DECREASED NUMBER OF
CC       HEMIDESMOSOMES. LAMININ-5 IS MISSING FROM THE BASEMENT MEMBRANE OF
CC       PATIENTS WITH THE GRAVIS FORM OF EPIDERMOLYSIS BULLOSA.
CC   -!- SIMILARITY: CONTAINS 2.5 LAMININ EGF-LIKE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L34155; AAA59483.1; -.
DR   EMBL; X85107; CAA59428.1; -.
DR   EMBL; X85108; CAA59429.1; -.
DR   HSSP; P02468; 1TLE.
DR   MIM; 600805; -.
DR   MIM; 226700; -.
DR   INTERPRO; IPR000561; -.
DR   INTERPRO; IPR001791; -.
DR   INTERPRO; IPR002049; -.
DR   PFAM; PF00053; laminin_EGF; 2.
DR   PFAM; PF00054; laminin_G; 2.
DR   PROSITE; PS00022; EGF_1; 1.
DR   PROSITE; PS01186; EGF_2; 1.
DR   PROSITE; PS01248; LAMININ_TYPE_EGF; 2.
KW   Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW   Laminin EGF-like domain; Cell adhesion; Repeat; Signal;
KW   Alternative splicing.
FT   SIGNAL       1     20       POTENTIAL.
FT   CHAIN       21   1713       LAMININ ALPHA-3 CHAIN.
FT   DOMAIN      46    201       DOMAIN III A.
FT   DOMAIN      67    185       2.5 X LAMININ EGF-LIKE REPEATS.
FT   DOMAIN      67    113       LAMININ EGF-LIKE 1.
FT   DOMAIN     114    166       LAMININ EGF-LIKE 2.
FT   DOMAIN     167    185       LAMININ EGF-LIKE 3 (INCOMPLETE).
FT   DOMAIN     202    793       DOMAIN II AND I (HEPTAT REPEATS).
FT   DOMAIN     794   1713       5 X LAMININ G-LIKE REPEATS (DOMAIN G).
FT   DOMAIN     794    970       LAMININ G-LIKE 1.
FT   DOMAIN     971   1139       LAMININ G-LIKE 2.
FT   DOMAIN    1140   1353       LAMININ G-LIKE 3.
FT   DOMAIN    1354   1529       LAMININ G-LIKE 4.
FT   DOMAIN    1530   1713       LAMININ G-LIKE 5.
FT   DOMAIN     231    327       COILED COIL (POTENTIAL).
FT   DOMAIN     396    548       COILED COIL (POTENTIAL).
FT   DOMAIN     594    621       COILED COIL (POTENTIAL).
FT   DOMAIN     702    765       COILED COIL (POTENTIAL).
FT   DOMAIN    1686   1713       COILED COIL (POTENTIAL).
FT   DISULFID    67     76       BY SIMILARITY.
FT   DISULFID    69     83       BY SIMILARITY.
FT   DISULFID    86     95       BY SIMILARITY.
FT   DISULFID    98    111       BY SIMILARITY.
FT   DISULFID   114    126       BY SIMILARITY.
FT   DISULFID   116    135       BY SIMILARITY.
FT   DISULFID   137    146       BY SIMILARITY.
FT   DISULFID   149    164       BY SIMILARITY.
FT   DISULFID   202    202       INTERCHAIN (PROBABLE).
FT   DISULFID   205    205       INTERCHAIN (PROBABLE).
FT   SITE       658    660       CELL ATTACHMENT SITE (POTENTIAL).
FT   CARBOHYD   542    542       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   645    645       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   745    745       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   882    882       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   964    964       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1108   1108       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1131   1131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1325   1325       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1477   1477       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD  1667   1667       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     1     45       MGWLWIFGAALGQCLGYSSQQQRVPFLQPPGQSQLQASYVE
FT                               FRPS -> KVSSYGGYLTYQAKSFALPGDMVLLEKKPDVQL
FT                               TGQHMSIIYEETNTPRPDRLHHGRVHVVEGNFRHASSRAPV
FT                               SREELMTVLSRLADVRIQGLYFTETQRLTLSEVGLEEASDT
FT                               GSGRIALAVEICACPPAYAGDSC (IN ISOFORM B).
FT   CONFLICT     5      5       W -> R (IN REF. 2).
FT   CONFLICT   123    125       ATG -> GMC (IN REF. 2).
FT   CONFLICT   481    481       M -> K (IN REF. 2).
FT   CONFLICT   754    754       R -> L (IN REF. 2).
FT   CONFLICT   969    969       E -> Q (IN REF. 2).
FT   CONFLICT  1052   1052       D -> A (IN REF. 2).
FT   CONFLICT  1184   1184       G -> A (IN REF. 2).
SQ   SEQUENCE  1713 AA;  189304 MW;  45EA9BE1017B60D3 CRC64;
```

Query Match 19.0%; Score 68.5; DB 1; Length 1713;
Best Local Similarity 35.5%; Pred. No. 12;
Matches 22; Conservative 12; Mismatches 21; Indels 7; Gaps 3;

```
Qy       8 AVDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIF 64
           |||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |::
Db     552 AVDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLL 607

Qy      65 DK 66
            ::
Db     608 NE 609
```

Search completed: February 3, 2001, 02:24:29
Job time: 641 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:17:41 ; Search time 118.74 Seconds
(without alignments)
40.029 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_66:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	350	97.0	352	1	QQCVL1	AL1 protein - toma
2	250	69.3	361	1	QQCVPT	AL1 protein - pota
3	237	65.7	358	2	S07594	hypothetical prote
4	230	63.7	362	1	JQ1887	AL1 protein - toma
5	223	61.8	359	2	S39211	gene C1 protein -
6	219	60.7	349	2	JQ2300	replicase - pepper
7	219	60.7	349	2	S31875	AL1 protein - pepp
8	216	59.8	359	2	S22593	hypothetical prote
9	213	59.0	351	2	JQ2327	AL1 protein - Indi
10	213	59.0	355	1	QQCVW1	AV1 protein - abut
11	212	58.7	357	1	QQCVC1	AL1 protein - toma
12	210	58.2	358	1	JQ1870	AL1 protein - toma
13	210	58.2	359	2	S39235	gene C1 protein -
14	210	58.2	385	2	S28360	AL1 protein - beet
15	203	56.2	360	2	S59885	replication-associ
16	142	39.3	131	2	S45059	AC1 protein (clone
17	117	32.4	347	1	QQCVS1	AL1 protein - squa
18	68.5	19.0	1713	2	A55347	adhesive ligand ep
19	68	18.8	587	2	JC1419	Fc gamma (IgG) rec
20	66.5	18.4	201	2	A81380	probable flagellar
21	64.5	17.9	299	2	B71967	probable peptidyl-
22	63.5	17.6	447	2	T12544	hypothetical prote
23	63.5	17.6	1610	2	A46227	voltage-dependent
24	63.5	17.6	1646	2	JH0422	voltage-dependent
25	63.5	17.6	2161	2	JH0564	calcium channel al
26	63.5	17.6	2181	2	A38198	calcium channel al
27	63.5	17.6	2203	2	T42742	voltage-dependent
28	62	17.2	295	2	D42452	C1 protein - tobac
29	61.5	17.0	481	2	A70091	probable phosphoes
30	61	16.9	1502	1	RGBYH1	CYC1/CYP3 transcri
31	60	16.6	432	2	D75348	gamma-glutamyl pho
32	60	16.6	447	2	S52437	CDP-diacylglycerol
33	59.5	16.5	136	2	T22240	hypothetical prote
34	59.5	16.5	299	2	G64541	cell binding facto
35	59.5	16.5	388	2	C69196	hypothetical prote
36	59.5	16.5	2137	2	T05244	hypothetical prote
37	59	16.3	316	2	C82085	conserved hypothet
38	59	16.3	397	2	B71078	probable NADH oxid
39	58.5	16.2	297	1	MNVNRV	nonstructural prot
40	58.5	16.2	335	2	JS0164	glyceraldehyde-3-p
41	58.5	16.2	589	2	D64125	hypothetical prote
42	58.5	16.2	1044	2	T43800	protein kinase bub
43	58.5	16.2	1070	1	A54600	1-phosphatidylinos
44	58.5	16.2	1265	2	T47626	structural mainten
45	58.5	16.2	1743	2	T15893	hypothetical prote

ALIGNMENTS

RESULT 1
QQCVL1
AL1 protein - tomato golden mosaic virus
C;Species: tomato golden mosaic virus
A;Note: host Nicotiana sp. (tobacco)
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 08-Apr-1994
C;Accession: A04170
R;Hamilton, W.D.O.; Stein, V.E.; Coutts, R.H.A.; Buck, K.W.
EMBO J. 3, 2197-2205, 1984
A;Title: Complete nucleotide sequence of the infectious cloned DNA components of toma
A;Reference number: A04163
A;Accession: A04170
A;Molecule type: DNA
A;Residues: 1-352 <HAM>
C;Comment: The genome consists of two circular, single-stranded DNA components, DNA A
.
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 97.0%; Score 350; DB 1; Length 352;
Best Local Similarity 97.1%; Pred. No. 1.2e-31;
Matches 68; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     111 TLVWGEFQVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 170

Qy      61 DRIFDKTPEP 70
           ||||||||||
Db     171 DRIFDKTPEP 180
```

RESULT 2
QQCVPT
AL1 protein - potato yellow mosaic virus (isolate Venezuela)
C;Species: potato yellow mosaic virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C;Accession: JU0364
R;Coutts, R.H.A.; Coffin, R.S.; Roberts, E.J.F.; Hamilton, W.D.O.
J. Gen. Virol. 72, 1515-1520, 1991
A;Title: The nucleotide sequence of the infectious cloned DNA components of potato ye
A;Reference number: JU0362; MUID:91311403
A;Accession: JU0364
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-361 <COU>
A;Cross-references: GB:D00940; NID:g222458; PIDN:BAA00782.1; PID:g222459
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 69.3%; Score 250; DB 1; Length 361;
Best Local Similarity 68.1%; Pred. No. 2e-20;
Matches 47; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

```
Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||   :|||||||| || |||||||||: :|| |::||:||:|||:|||:|||: ||
Db    110 TIEWGLFQIDGRSARGGQQTVNDAAAEALNSGTKEAAMKIIKEKLPEKFLFQYHNLSCNL 169

Qy     61 DRIFDKTPE 69
          |||| | ||
Db    170 DRIFMKAPE 178
```

RESULT 3
S07594
hypothetical protein, 40.4K - cassava latent virus (Nigerian isolate)
C;Species: cassava latent virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Sep-1999
C;Accession: S07594
R;Morris, B.; Coates, L.; Lowe, S.; Richardson, K.; Eddy, P.
Nucleic Acids Res. 18, 197-198, 1990
A;Title: Nucleotide sequence of the infectious cloned DNA components of African cassava
A;Reference number: S07590; MUID:90174930
A;Accession: S07594
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <MOR>
A;Cross-references: EMBL:X17095; NID:g59371; PIDN:CAA34953.1; PID:g59376
C;Genetics:
A;Map position: segment DNA1
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 65.7%; Score 237; DB 2; Length 358;
Best Local Similarity 61.4%; Pred. No. 5.7e-19;
Matches 43; Conservative 13; Mismatches 14; Indels 0; Gaps 0;

```
Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  :|||||||| |::||| |:||: || ||| :||| :|: :: ||||||||||
Db    109 TVEWGQFQIDGRSARGGQQSANDAYAKALNSGSKSEALNVIRELVPKDFVLQFHNLNSNL 168

Qy     61 DRIFDKTPEP 70
          |||| : | |
Db    169 DRIFQEPPAP 178
```

RESULT 4
JQ1887
AL1 protein - tomato yellow leaf curl virus (strain Australia)
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1887
R;Dry, I.B.; Rigden, J.E.; Krake, L.R.; Mullineaux, P.M.; Rezaian, M.A.
J. Gen. Virol. 74, 147-151, 1993
A;Title: Nucleotide sequence and genome organization of tomato leaf curl geminivirus.
A;Reference number: JQ1885; MUID:93139778
A;Accession: JQ1887
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-362 <DRY>
A;Cross-references: GB:S53251
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 63.7%; Score 230; DB 1; Length 362;
Best Local Similarity 54.1%; Pred. No. 3.5e-18;
Matches 46; Conservative 9; Mismatches 14; Indels 16; Gaps 1;

```
Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          || |||  :|||||||| |::||| |:|||  || ||| ::||  |: |: |||||||||
Db    110 TLEWGEFQIDGRSARGGQQSANDAYAQALNTGSKSEALNVLRELAPKDYVLQFHNLNSNL 169

Qy     61 DRI----------------FDKTPE 69
          |||                ||: ||
Db    170 DRIFTPPLEVYVSPFLSSSFDRVPE 194
```

RESULT 5
S39211
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Sep-1999
C;Accession: S39211
R;Noris, E.; Hidalgo, E.; Accotto, G.; Moriones, E.
submitted to the EMBL Data Library, August 1993
A;Description: High similarity among the tomato yellow leaf curl virus isolates from
A;Reference number: S39209
A;Accession: S39211
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <NOR>
A;Cross-references: EMBL:Z25751; NID:g433655; PIDN:CAA81026.1; PID:g433658
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 61.8%; Score 223; DB 2; Length 359;
Best Local Similarity 59.4%; Pred. No. 2.1e-17;
Matches 41; Conservative 9; Mismatches 19; Indels 0; Gaps 0;

```
Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || ||| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSEALDVIKELAPRDYILHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          |:|   | |
Db    171 RVFQVPPAP 179
```

RESULT 6
JQ2300
replicase - pepper huasteco virus (component A)
N;Alternate names: ORF AL1 protein
C;Species: pepper huasteco virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Sep-1999
C;Accession: JQ2300
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Herrera-Estrella, L.; Rivera-Bustamante,
J. Gen. Virol. 74, 2225-2231, 1993
A;Title: Complete nucleotide of pepper huasteco virus: Analysis and comparison with b
A;Reference number: JQ2299; MUID:94015007
A;Accession: JQ2300
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: GB:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
C;Superfamily: tomato golden mosaic virus AL1 protein

Query Match 60.7%; Score 219; DB 2; Length 349;
Best Local Similarity 58.6%; Pred. No. 5.7e-17;
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||  :|||||||| |::||  |:|||::| |||||||:|: |: :  ||||: ||
Db    110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy     61 DRIFDKTPEP 70
          :|||   |||
Db    170 NRIFQTPPEP 179
```

RESULT 7
S31875
AL1 protein - pepper rizado amarillo virus
C;Species: pepper rizado amarillo virus

```
C;Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C;Accession: S31875
R;Torres-Pacheco, I.; Garzon-Tiznado, J.A.; Jimenez, B.; Herrera-Estrella, L.; Rivera-Bu
submitted to the EMBL Data Library, February 1993
A;Description: Complete nucleotide sequence of pepper huasteco virus: analysis and compa
A;Reference number: S31872
A;Accession: S31875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-349 <TOR>
A;Cross-references: EMBL:X70418; NID:g61023; PIDN:CAA49856.1; PID:g61025
A;Note: the source is designated as pepper huasteco virus
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             60.7%;  Score 219;  DB 2;  Length 349;
  Best Local Similarity   58.6%;  Pred. No. 5.7e-17;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: |||  :|||||||| |::|| |:|||::| |||||||:|: |: :  ||||: ||
Db     110 TVEWGEFQIDGRSARGGQQSANDTYAKALNSASAEEALQIIKEEQPQHFFLQFHNIVSNA 169

Qy      61 DRIFDKTPEP 70
           :|||   |||
Db     170 NRIFQTPPEP 179


RESULT   8
S22593
hypothetical protein C4 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Sep-1999
C;Accession: S22593
R;Kheyr-Pour, A.; Bendahmane, M.; Matzeit, V.; Accotto, G.P.; Crespi, S.; Gronenborn, B.
Nucleic Acids Res. 19, 6763-6769, 1991
A;Title: Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monoparti
A;Reference number: S22588; MUID:92107660
A;Accession: S22593
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-359 <KHE>
A;Cross-references: EMBL:X61153; NID:g62211; PIDN:CAA43466.1; PID:g62217
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.8%;  Score 216;  DB 2;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 1.3e-16;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   9
JQ2327
AL1 protein - Indian cassava mosaic virus
N;Alternate names: replication-associated protein
C;Species: Indian cassava mosaic virus
C;Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2327; S35883
R;Hong, Y.G.; Robinson, D.J.; Harrison, B.D.
J. Gen. Virol. 74, 2437-2443, 1993
A;Title: Nucleotide sequence evidence for the occurrence of three distinct whitefly-tran
A;Reference number: JQ2326; MUID:94065670
A;Accession: JQ2327
A;Molecule type: DNA
A;Residues: 1-351 <HON>
A;Cross-references: EMBL:Z24758; NID:g395351; PIDN:CAA80891.1; PID:g584046
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.0%;  Score 213;  DB 2;  Length 351;
  Best Local Similarity   61.2%;  Pred. No. 2.7e-16;
  Matches   41;  Conservative    9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           ||   :|||||||| |::||| | |||: || |||:|:||  |  ||  ||:::||||||
Db     113 WGTFQIDGRSARGGQQSANDAYAAALNSGSKSEALKILRELAPRDYLRDFHHISSNLDRI 172

Qy      64 FDKTPEP 70
           | | | |
Db     173 FTKPPPP 179


RESULT  10
QQCVW1
AV1 protein - abutilon mosaic virus (isolate West India)
C;Species: abutilon mosaic virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 08-Apr-1994
C;Accession: A36214
R;Frischmuth, T.; Zimmat, G.; Jeske, H.
Virology 178, 461-468, 1990
A;Title: The nucleotide sequence of the abutilon mosaic virus reveals prokaryotic as
A;Reference number: A36214; MUID:91020984
A;Accession: A36214
A;Molecule type: DNA
A;Residues: 1-355 <FRI>
A;Cross-references: EMBL:X15983
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             59.0%;  Score 213;  DB 1;  Length 355;
  Best Local Similarity   57.1%;  Pred. No. 2.7e-16;
  Matches   40;  Conservative   12;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |  ||| :|||||||| ||:||: |:|||    : || |::|: |: |: | ||: |||
Db     110 TAEWGEFQIDGRSARGGQQTANDSYAKALNAGDVQSALNILKEEQPKDYVLQNHNIRSNL 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 ERIFAKAPEP 179


RESULT  11
QQCVC1
AL1 protein - tomato yellow leaf curl virus
N;Alternate names: C1 protein
C;Species: tomato yellow leaf curl virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 23-Jul-1999
C;Accession: D40779
R;Navot, N.; Pichersky, E.; Zeidan, M.; Zamir, D.; Czosnek, H.
Virology 185, 151-161, 1991
A;Title: Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a sin
A;Reference number: A40779; MUID:92024070
A;Accession: D40779
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-357 <NAV>
A;Cross-references: GB:X15656; NID:g62204; PIDN:CAA33688.1; PID:g62207
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.7%;  Score 212;  DB 1;  Length 357;
  Best Local Similarity   65.6%;  Pred. No. 3.6e-16;
  Matches   40;  Conservative   11;  Mismatches   10;  Indels    0;  Gaps    0;
```

```
Qy      4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :| : :|||||||| |::||| ||||| || ||| |::|| |: |: ||||||||||||
Db    111 FGVSQIDGRSARGGQQSANDAYAEALNSGSKSEALNILKEKAPKDYILQFHNLSSNLDRI 170

Qy     64 F 64
          |
Db    171 F 171


RESULT  12
JQ1870
AL1 protein - tomato mottle virus (isolate Florida)
C;Species: tomato mottle virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 07-May-1999
C;Accession: JQ1870
R;Abouzid, A.M.; Polston, J.E.; Hiebert, E.
J. Gen. Virol. 73, 3225-3229, 1992
A;Title: The nucleotide sequence of tomato mottle virus, a new geminivirus isolated from
A;Reference number: JQ1869; MUID:93107858
A;Accession: JQ1870
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-358 <ABO>
A;Cross-references: GB:L14460
C;Genetics:
A;Map position: segment A
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 210;  DB 1;  Length 358;
  Best Local Similarity   54.3%;  Pred. No. 6e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  :|||||||| |::||: |:|||| | : || ::||: |: :: | ||: |||
Db    107 TIEWGDFQIDGRSARGGQQSANDSYAKALNAGSVQSALAVLREEQPKDFVLQNHNIRSNL 166

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    167 ERIFAKAPEP 176


RESULT  13
S39235
gene C1 protein - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
C;Accession: S39235
R;Crespi, S.; Noris, E.; Vaira, A.; Bosco, D.; Accotto, G.
submitted to the EMBL Data Library, December 1993
A;Description: A cloned DNA from a TYLCV isolate from Sicily showing low infectivity.
A;Reference number: S39233
A;Accession: S39235
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <CRE>
A;Cross-references: EMBL:Z28390; NID:g1041671; PID:g1334964
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 210;  DB 2;  Length 359;
  Best Local Similarity   55.1%;  Pred. No. 6e-16;
  Matches   38;  Conservative   12;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || ||| :|::  |  |:  |||::||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINARSKSEALDVIKQLAPRDYVLHFHNISSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT  14
S28360
AL1 protein - beet curly top virus
C;Species: beet curly top virus
C;Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 20-Sep-1999
C;Accession: S28360
R;Stanley, J.; Markham, P.G.; Callis, R.J.; Pinner, M.S.
EMBO J. 5, 1761-1767, 1986
A;Title: The nucleotide sequence of an infectious clone of the geminivirus beet curly
A;Reference number: S28360
A;Accession: S28360
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-385 <STA>
A;Cross-references: GB:M24597; EMBL:X04144; NID:g210678; PIDN:AAA42751.1; PID:g210679
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             58.2%;  Score 210;  DB 2;  Length 385;
  Best Local Similarity   54.3%;  Pred. No. 6.5e-16;
  Matches   38;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: |||  :|||||||| ||:||: |:||||:| ::||||::|: |: |  | ||| :|
Db    137 TIEWGEFQIDGRSARGGQQTANDSYAKALNATSLDQALQILKEEQPKDYFLQHHNLLNNA 196

Qy     61 DRIFDKTPEP 70
          :|| : |:|
Db    197 QKIFQRPPDP 206


RESULT  15
S59885
replication-associated protein C1 - tomato yellow leaf curl virus
C;Species: tomato yellow leaf curl virus
C;Date: 14-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 20-Sep-1999
C;Accession: S59885
R;Hong, Y.; Harrison, B.D.
submitted to the EMBL Data Library, February 1995
A;Description: Nucleotide sequences from tomato leaf curl viruses from different coun
d geminiviruses.
A;Reference number: S58346
A;Accession: S59885
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <HON>
A;Cross-references: EMBL:Z48182; NID:g944838; PIDN:CAA88229.1; PID:g974211
C;Superfamily: tomato golden mosaic virus AL1 protein


  Query Match             56.2%;  Score 203;  DB 2;  Length 360;
  Best Local Similarity   59.1%;  Pred. No. 3.7e-15;
  Matches   39;  Conservative   12;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
          :|   :|||||||| |::||| |||:|: ||  || |:||| |: :: ||||||:|||||
Db    113 FGVFQIDGRSARGGQQSANDAYAEAINSGSKAAALDILREKAPKDFVLQFHNLNANLDRI 172

Qy     64 FDKTPE 69
          | : |
Db    173 FTPSAE 178



Search completed: February  3, 2001, 02:17:41
Job time: 2213 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:19:48 ; Search time 109.09 Seconds
(without alignments)
11.523 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	216	59.8	359	3	US-08-809-103B-2	Sequence 2, Appli
2	216	59.8	359	3	US-08-809-103B-4	Sequence 4, Appli
3	216	59.8	359	3	US-08-809-103B-6	Sequence 6, Appli
4	216	59.8	359	3	US-08-809-103B-8	Sequence 8, Appli
5	68.5	19.0	1713	3	US-08-600-982-24	Sequence 24, Appl
6	68.5	19.0	1713	4	PCT-US94-10261A-24	Sequence 24, Appl
7	63.5	17.6	2161	1	US-07-745-206A-2	Sequence 2, Appli
8	63.5	17.6	2161	1	US-08-455-543A-49	Sequence 49, Appl
9	63.5	17.6	2161	1	US-08-455-543A-51	Sequence 51, Appl
10	63.5	17.6	2161	2	US-08-223-305C-49	Sequence 49, Appl
11	63.5	17.6	2161	2	US-08-223-305C-51	Sequence 51, Appl
12	63.5	17.6	2161	2	US-08-311-363-2	Sequence 2, Appli
13	60	16.6	446	2	US-08-672-814D-11	Sequence 11, Appl
14	58	16.1	171	2	US-08-609-049A-22	Sequence 22, Appl
15	58	16.1	844	3	US-09-029-267-20	Sequence 20, Appl
16	57.5	15.9	334	5	5290690-11	Patent No. 5290690
17	57.5	15.9	439	1	US-07-637-870-9	Sequence 9, Appli
18	57.5	15.9	439	1	US-07-637-399-6	Sequence 6, Appli
19	57.5	15.9	439	1	US-08-112-703-6	Sequence 6, Appli
20	57	15.8	165	4	PCT-US95-10075-4	Sequence 4, Appli
21	56	15.5	454	3	US-08-446-100-31	Sequence 31, Appl
22	54.5	15.1	103	1	US-08-552-142A-13	Sequence 13, Appl
23	54.5	15.1	497	2	US-08-570-311-2	Sequence 2, Appli
24	54.5	15.1	497	2	US-08-353-485-2	Sequence 2, Appli
25	54.5	15.1	2628	2	US-08-570-311-14	Sequence 14, Appl
26	54	15.0	454	3	US-08-446-100-30	Sequence 30, Appl
27	54	15.0	602	2	US-08-419-652-6	Sequence 6, Appli
28	54	15.0	771	1	US-07-923-976-6	Sequence 6, Appli
29	54	15.0	783	5	5422248-2	Patent No. 5422248
30	54	15.0	836	1	US-07-923-976-4	Sequence 4, Appli
31	54	15.0	863	1	US-07-923-976-8	Sequence 8, Appli
32	53.5	14.8	439	1	US-07-762-681A-1	Sequence 1, Appli
33	53	14.7	81	3	US-08-995-159-6	Sequence 6, Appli
34	53	14.7	266	2	US-07-857-224B-40	Sequence 40, Appl
35	53	14.7	312	1	US-08-414-625-2	Sequence 2, Appli
36	53	14.7	454	3	US-08-446-100-26	Sequence 26, Appl
37	53	14.7	454	3	US-08-446-100-27	Sequence 27, Appl
38	53	14.7	454	3	US-08-446-100-28	Sequence 28, Appl
39	53	14.7	454	3	US-08-446-100-29	Sequence 29, Appl
40	53	14.7	548	2	US-08-871-266B-17	Sequence 17, Appl
41	53	14.7	548	2	US-09-018-864A-17	Sequence 17, Appl
42	53	14.7	548	3	US-08-871-267B-23	Sequence 23, Appl
43	53	14.7	783	5	5231168-2	Patent No. 5231168
44	53	14.7	1832	3	US-09-335-409-4	Sequence 4, Appli
45	52.5	14.5	589	3	US-08-942-008-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-08-809-103B-2
; Sequence 2, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-2


  Query Match 59.8%;  Score 216;  DB 3;  Length 359;
  Best Local Similarity  56.5%;  Pred. No. 3.1e-20;
  Matches  39;  Conservative  11;  Mismatches  19;  Indels  0;  Gaps  0;
```

```
Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   2
US-08-809-103B-4
; Sequence 4, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-4


  Query Match             59.8%;  Score 216;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   3
US-08-809-103B-6
; Sequence 6, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,103B
;      FILING DATE:  17-MAR-1997
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 94.11040
;      FILING DATE:  15-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/FR95/01192
;      FILING DATE:  15-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  PATCH, Andrew J.
;      REGISTRATION NUMBER:  32,925
;      REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 521-2297
;      TELEFAX:  (703) 685-0573
;      TELEX:  248425 EMBON
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  359 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-809-103B-6


  Query Match             59.8%;  Score 216;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
          | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db    111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy     62 RIFDKTPEP 70
          ::|   | |
Db    171 KVFQVPPAP 179


RESULT   4
US-08-809-103B-8
; Sequence 8, Application US/08809103B
; Patent No. 6133505
;  GENERAL INFORMATION:
;    APPLICANT:  GRONENBORN, Bruno
;    TITLE OF INVENTION:  PHYTOPATHOGENIC DNA VIRUS RESISTANT
;    TITLE OF INVENTION:  TRANSGENIC PLANTS AND SEEDS AND METHODS FOR OBTAINING SAME
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  YOUNG & THOMPSON
;      STREET:  745 South 23rd Street
;      CITY:  Arlington
;      STATE:  Virginia
```

```
;       COUNTRY:  U.S.A.
;       ZIP:  22202
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/809,103B
;       FILING DATE:  17-MAR-1997
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 94.11040
;       FILING DATE:  15-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR95/01192
;       FILING DATE:  15-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  PATCH, Andrew J.
;       REGISTRATION NUMBER:  32,925
;       REFERENCE/DOCKET NUMBER:  USB94AL CNR TOM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 521-2297
;       TELEFAX:  (703) 685-0573
;       TELEX:  248425 EMBON
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  359 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-809-103B-8


  Query Match             59.8%;  Score 216;  DB 3;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-20;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :||||||||| ||:||| |:|:|| || :|| :|:|   |  |:  |||:|||||
Db     111 LEWGTFQIDGRSARGGQQTANDAYAKAINAGSKSQALDVIKELAPRDYVLHFHNINSNLD 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 KVFQVPPAP 179


RESULT   5
US-08-600-982-24
; Sequence 24, Application US/08600982
; Patent No. 6120991
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/600,982
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence
;      DESCRIPTION:  of FIGURES 15A-15F, and as shown also in FIGURES
;      DESCRIPTION:  19A-19R
US-08-600-982-24


  Query Match             19.0%;  Score 68.5;  DB 3;  Length 1713;
  Best Local Similarity   35.5%;  Pred. No. 3.5;
  Matches   22;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       8 AVDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIF 64
             |||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |::
Db     552 AVDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLL 607

Qy      65 DK 66
           ::
Db     608 NE 609


RESULT   6
PCT-US94-10261A-24
; Sequence 24, Application PC/TUS9410261A
;  GENERAL INFORMATION:
;    APPLICANT:  Carter, William G.
;    APPLICANT:  Gil, Susanna A.
;    APPLICANT:  Ryan, Maureen C.
;    TITLE OF INVENTION:  Epiligrin, an Epithelial Ligand for
;    TITLE OF INVENTION:  Integrins
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christensen, O'Connor, Johnson, and Kindness
;      STREET:  1420 Fifth Avenue
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98101-8100
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10261A
;      FILING DATE:  02-SEP-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Shelton, Dennis K.
;      REGISTRATION NUMBER:  26,997
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 682-8100
;      TELEFAX:  (206) 224-0779
;  INFORMATION FOR SEQ ID NO:  24:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1713 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;      DESCRIPTION:  E170 protein as translated from sequence of
;      DESCRIPTION:  FIGURES 15A-15F, and as shown also in FIGURES 19A-19R
PCT-US94-10261A-24
```

```
  Query Match             19.0%;  Score 68.5;  DB 4;  Length 1713;
  Best Local Similarity   35.5%;  Pred. No. 3.5;
  Matches   22;  Conservative   12;  Mismatches   21;  Indels    7;  Gaps    3;

Qy       8 AVDGRSARGGCQTSNDAAAEALN--ASSKEEALQ-IIREKIPEKYLFQFHNLNSNLDRIF 64
           |||  :|      :  || :| |  ||: | ||| :|:| :| |       |:|| |::
Db     552 AVDAATAYENILNAIKAAEDAANRAASASESALQTVIKEDLPRK----AKTLSSNSDKLL 607

Qy      65 DK 66
           ::
Db     608 NE 609


RESULT   7
US-07-745-206A-2
; Sequence 2, Application US/07745206A
; Patent No. 5429921
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Feldman, Daniel
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fitch, Even, Tabin & Flannery
;      STREET:  135 S. LaSalle
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  U.S.A.
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/745,206A
;      FILING DATE:  19910815
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Feder, Scott B
;      REFERENCE/DOCKET NUMBER:  51504
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-372-7842
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-745-206A-2


  Query Match             17.6%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 21;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEAAVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                            | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   8
US-08-455-543A-49
; Sequence 49, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-49


  Query Match             17.6%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 21;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;
```

```
Qy       4 WGEAAVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT   9
US-08-455-543A-51
; Sequence 51, Application US/08455543A
; Patent No. 5792846
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/455,543A
;      FILING DATE:  May 31, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/223,305
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-52517
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-455-543A-51


  Query Match             17.6%;  Score 63.5;  DB 1;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 21;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEAAVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  10
US-08-223-305C-49
; Sequence 49, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
```

```
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  49:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-49


  Query Match             17.6%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 21;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEAAVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  11
US-08-223-305C-51
; Sequence 51, Application US/08223305C
; Patent No. 5851824
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  HUMAN CALCIUM CHANNEL COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/223,305C
;      FILING DATE:  April 4, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/868,354
;      FILING DATE:  April 10, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/620,250
;      FILING DATE:  30-NOV-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/482,384
;      FILING DATE:  20-FEB-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/603,751
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US89/01408
;      FILING DATE:  04-APR-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/176,899
;      FILING DATE:  04-APR-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  52516  (P519739)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2161 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-08-223-305C-51


  Query Match             17.6%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 21;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEAAVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  12
US-08-311-363-2
; Sequence 2, Application US/08311363
; Patent No. 5876958
;  GENERAL INFORMATION:
;    APPLICANT:  Harpold, Michael
;    APPLICANT:  Ellis, Steven
;    APPLICANT:  Williams, Mark
;    APPLICANT:  Feldman, Daniel
;    APPLICANT:  McCue, Ann
;    APPLICANT:  Brenner, Robert
;    TITLE OF INVENTION:  Human Calcium Channel Compositions and
;    TITLE OF INVENTION:  Methods
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/311,363
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/745,206
;      FILING DATE:  15-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6362-51506
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  (619)238-0999
;      TELEFAX:  (619)238-0062
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2161 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-311-363-2


  Query Match             17.6%;  Score 63.5;  DB 2;  Length 2161;
  Best Local Similarity   28.4%;  Pred. No. 21;
  Matches   21;  Conservative    5;  Mismatches   19;  Indels   29;  Gaps    1;

Qy       4 WGEAAVDGRSARGGCQTS----------------------------NDAAAEALNASSK 34
           |     ||  | ||  :|                              | | ||:|| : |
Db     707 WNAVMYDGIMAYGGPSSSGMIVCIYFIILFICGNYILLNVFLAIAVDNLADAESLNTAQK 766

Qy      35 EEALQIIREKIPEK 48
           ||| :  |:||  |
Db     767 EEAEEKERKKIARK 780


RESULT  13
US-08-672-814D-11
; Sequence 11, Application US/08672814D
; Patent No. 5952480
;  GENERAL INFORMATION:
;     APPLICANT:  Leung, David W.
;     APPLICANT:  Weeks, Reitha
;     TITLE OF INVENTION:  MAMMALIAN
;     NUMBER OF SEQUENCES:  13
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cell Therapeutics, Inc.
;       STREET:  201 Elliott Avenue West
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  U.S.A.
;       ZIP:  98119
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" disk, 1.44Mb, double side, high density
;       COMPUTER:  PC Clone (486 microprocessor)
;       OPERATING SYSTEM:  MS-DOS Version 6.1, Windows NT
;       SOFTWARE:  WORD 6.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/672,814D
;       FILING DATE:  28-Jun-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Faciszewski, Stephen
;       REGISTRATION NUMBER:  36,131
;       REFERENCE/DOCKET NUMBER:  1803
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206)282-7100
;       TELEFAX:  (206)284-6206
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  446
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  no
;     ANTI-SENSE:  no
;     FRAGMENT TYPE:
;     ORIGINAL SOURCE:
;       ORGANISM:  Drosophila
;       STRAIN:
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:
;       HAPLOTYPE:
;       TISSUE TYPE:
;       CELL TYPE:
;       CELL LINE:
;       ORGANELLE:
;     FEATURE:
;       NAME/KEY:  CDP-diacylglycerol synthase (CDS)
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
US-08-672-814D-11


  Query Match             16.6%;  Score 60;  DB 2;  Length 446;
  Best Local Similarity   34.0%;  Pred. No. 7.8;
  Matches   17;  Conservative    9;  Mismatches   22;  Indels    2;  Gaps    1;

Qy      20 TSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRIFDKTPE 69
           : :||| :  :|:   : :    ||||||:        |  || :  |||||
Db      20 SGSDAANKRNSAADSSDHVDSEEEKIPEEKFVD--ELAKNLPQGTDKTPE 67


RESULT  14
US-08-609-049A-22
; Sequence 22, Application US/08609049A
; Patent No. 5948664
;  GENERAL INFORMATION:
;     APPLICANT:  Williams, Lewis T.
;     APPLICANT:  Molz, Lisa
;     APPLICANT:  Chen, Yen-Wen
;     TITLE OF INVENTION:  No. 5948664el PI 3-Kinase Polypeptides
;     NUMBER OF SEQUENCES:  32
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Townsend and Townsend and Crew LLP
;       STREET:  Two Embarcadero Center, 8th Floor
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94111-3834
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/609,049A
;       FILING DATE:  29-FEB-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Dow, Karen B.
;       REGISTRATION NUMBER:  29,684
;       REFERENCE/DOCKET NUMBER:  2307K-063700US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-326-2400
;       TELEFAX:  415-326-2422
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  171 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-609-049A-22


  Query Match             16.1%;  Score 58;  DB 2;  Length 171;
  Best Local Similarity   24.7%;  Pred. No. 4.1;
  Matches   22;  Conservative   11;  Mismatches   20;  Indels   36;  Gaps    3;

Qy       2 LVWGEAAVDGRSARGGC----------------------QTSNDAAAEALNASSKEEA 37
             |:| || :| |    ||                        :|| ||| |  | :|:
Db      24 LIWKEAGLDIRMLPYGCLATGDRSGLIEVVSTSETIADIQLNSSNVAAAAA--AFNKDAL 81

Qy      38 LQIIREKIPEKYLFQFHNLNSNLDRIFDK 66
           |  ::|          :|   :|||  ::
Db      82 LNWLKE----------YNSGDDLDRAIEE 100
```

```
RESULT 15
US-09-029-267-20
; Sequence 20, Application US/09029267
; Patent No. 6107057
;  GENERAL INFORMATION:
;    APPLICANT:  Crawford, Kenneth
;    APPLICANT:  Zaror, Isabel
;    APPLICANT:  Innis, Michael
;    TITLE OF INVENTION:  Pichia Secretory Leader for Protein
;    TITLE OF INVENTION:  Expression
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/029,267
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Chung, Ling-Fong
;      REGISTRATION NUMBER:  36,482
;      REFERENCE/DOCKET NUMBER:  1165.100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 601-2704
;      TELEFAX:  (510) 655-3542
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  844 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-029-267-20


  Query Match             16.1%;  Score 58;  DB 3;  Length 844;
  Best Local Similarity   27.5%;  Pred. No. 32;
  Matches   11;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      24 AAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
           |  :  |: :: |:||::  :  ||  : |  | :|: :|::
Db     714 AHSTTLSVATYEQALEMSNQYAPEHLILQIENASSYVDQV 753


Search completed: February  3, 2001, 02:19:49
Job time: 1711 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 3, 2001, 02:15:29 ; Search time 144.12 Seconds
(without alignments)
16.608 Million cell updates/sec

Title: US-09-289-346A-8
Perfect score: 361
Sequence: 1 TLVWGEAAVDGRSARGGCQT.....FQFHNLNSNLDRIFDKTPEP 70

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS6/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS6/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	216	59.8	353	18	W34338	Bean golden mosaic
2	216	59.8	353	18	W34332	Bean golden mosaic
3	216	59.8	353	18	W34333	Bean golden mosaic
4	216	59.8	353	18	W34334	Bean golden mosaic
5	216	59.8	353	18	W34335	Bean golden mosaic
6	216	59.8	359	17	R88870	Sardinian tomato y
7	216	59.8	359	17	R88871	Sardinian tomato y
8	216	59.8	359	17	R88872	Sardinian tomato y
9	214	59.3	353	8	P70407	ORF 4 gene product
10	214	59.3	361	18	W34336	Tomato mottle viru
11	214	59.3	361	18	W34324	Tomato mottle viru
12	214	59.3	361	18	W34325	Tomato mottle viru
13	214	59.3	361	18	W34326	Tomato mottle viru
14	212	58.7	357	18	W34329	Tomato yellow leaf
15	212	58.7	357	18	W34330	Tomato yellow leaf
16	212	58.7	357	18	W34331	Tomato yellow leaf
17	207.5	57.5	361	8	P70562	Product of ORF 4 f
18	204	56.5	357	18	W34337	Tomato yellow leaf
19	202	56.0	362	19	W56495	Tobacco leaf curli
20	68.5	19.0	1713	16	R70148	Deduced sequence o
21	66.5	18.4	512	19	W68473	HIV-1 strain YBF30
22	65	18.0	292	21	Y92317	Geminivirus Rep C1
23	65	18.0	335	21	Y92318	Mastrevirus Rep in
24	64.5	17.9	299	18	W55450	H. pylori ORF 02ae
25	63.5	17.6	447	21	Y87089	Human secreted pro
26	63.5	17.6	447	21	Y66744	Membrane-bound pro
27	63.5	17.6	447	21	Y50947	Human adult aorta
28	63.5	17.6	456	21	Y87190	Human secreted pro
29	63.5	17.6	619	13	R27651	Human calcium chan
30	63.5	17.6	2161	14	R33545	Sequence of the al
31	63.5	17.6	2161	16	R71001	Human neuronal cal
32	63.5	17.6	2161	16	R71002	Human neuronal cal
33	63.5	17.6	2161	19	W63137	Human calcium chan
34	63.5	17.6	2161	19	W63149	Human calcium chan
35	60.5	16.8	204	20	W89814	Protein encoded by
36	60.5	16.8	299	20	W89829	Protein encoded by
37	60.5	16.8	299	20	W89849	Antigen from clust
38	59.5	16.5	299	19	W98322	H. pylori GHPO 136
39	59.5	16.5	299	19	W73034	Helicobacter pylor
40	58.5	16.2	297	8	P70318	Protein M1 encoded
41	58	16.1	844	18	W09773	Histidinol dehydro
42	57.5	15.9	334	11	R03211	Amino acid sequenc
43	57	15.8	165	17	R91063	Turkey prepro-VIP.
44	56.5	15.7	249	20	W89822	Protein encoded by
45	56.5	15.7	269	20	W89860	Antigen 1 from clu

ALIGNMENTS

RESULT 1
W34338
ID W34338 standard; Protein; 353 AA.
XX
AC W34338;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 protein.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93314.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus

XX
PS Example 5; Page 100-102; 132pp; English.
XX
CC This sequence comprises the bean golden mosaic virus (BGMV) C1
CC protein that is required for replication. The invention involves
CC production of transgenic plants containing DNA comprising C1 or AC1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from BGMV, tomato mottle virus or tomato yellow leaf
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC that have mutations in the highly conserved DNA-nicking domain
CC and/or the NTP-binding domains.
XX
SQ Sequence 353 AA;

Query Match 59.8%; Score 216; DB 18; Length 353;
Best Local Similarity 58.6%; Pred. No. 3e-19;
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179
```

RESULT 2
W34332
ID W34332 standard; Protein; 353 AA.
XX
AC W34332;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA190 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93290.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 103-105; 132pp; English.
XX
CC This protein comprises a control mutant of the bean golden mosaic
CC virus (BGMV) C1 protein (see W34338) that is required for
CC replication. It is encoded by mutated C1 open reading frame
CC BGAC190 (see T93290). The invention involves production of
CC transgenic plants containing DNA comprising geminivirus C1 or AC1
CC wild-type or mutant sequences that negatively interfere in trans
CC with geminiviral replication during infection. Such transgenic
CC plants are resistant to viral infection. The AC1/C1 genes are
CC especially from BGMV, tomato mottle virus or tomato yellow leaf
CC curl virus (see T93282-93) and encode polypeptides (see W34324-35)
CC that have mutations in the highly conserved DNA-nicking and/or the
CC NTP-binding domains.
XX
SQ Sequence 353 AA;

Query Match 59.8%; Score 216; DB 18; Length 353;
Best Local Similarity 58.6%; Pred. No. 3e-19;
Matches 41; Conservative 13; Mismatches 16; Indels 0; Gaps 0;

```
Qy      1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
          |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db    110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy     61 DRIFDKTPEP 70
          :||| | |||
Db    170 erifvkvpep 179
```

RESULT 3
W34333
ID W34333 standard; Protein; 353 AA.
XX
AC W34333;
XX
DT 27-APR-1998 (first entry)
XX
DE Bean golden mosaic geminivirus C1 BGA221 mutant.
XX
KW Geminivirus; BGMV; C1 gene; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Bean golden mosaic virus type II isolate Guatemala.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93291.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 5; Page 107-109; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the bean
CC golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC for replication. It is encoded by mutated C1 open reading frame
CC BGAC221 (see T93291) and carries a mutation in the NTP-binding
CC domain. The invention involves production of transgenic plants
CC containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC sequences that negatively interfere in trans with geminiviral
CC replication during infection. Such transgenic plants are resistant
CC to viral infection. The AC1/C1 genes are especially from BGMV,
CC tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC and encode polypeptides (see W34324-35) that have mutations in the

```
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             59.8%;  Score 216;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 3e-19;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   4
W34334
ID   W34334 standard; Protein; 353 AA.
XX
AC   W34334;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA228 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93292.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 111-112; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC228 (see T93292) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             59.8%;  Score 216;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 3e-19;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   5
W34335
ID   W34335 standard; Protein; 353 AA.
XX
AC   W34335;
XX
DT   27-APR-1998  (first entry)
XX
DE   Bean golden mosaic geminivirus C1 BGA262 mutant.
XX
KW   Geminivirus; BGMV; C1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Bean golden mosaic virus type II isolate Guatemala.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93293.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 5; Page 115-116; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant of the bean
CC   golden mosaic virus (BGMV) C1 protein (see W34338) that is required
CC   for replication.  It is encoded by mutated C1 open reading frame
CC   BGAC262 (see T93293) and carries a mutation in the NTP-binding
CC   domain.  The invention involves production of transgenic plants
CC   containing DNA comprising geminivirus C1 or AC1 wild-type or mutant
CC   sequences that negatively interfere in trans with geminiviral
CC   replication during infection.  Such transgenic plants are resistant
CC   to viral infection.  The AC1/C1 genes are especially from BGMV,
CC   tomato mottle virus or tomato yellow leaf curl virus (see T93282-93)
CC   and encode polypeptides (see W34324-35) that have mutations in the
CC   highly conserved DNA-nicking and/or the NTP-binding domains.
XX
SQ   Sequence   353 AA;


  Query Match             59.8%;  Score 216;  DB 18;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 3e-19;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqhhnirsnl 169
```

```
Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifvkvpep 179


RESULT   6
R88870
ID   R88870 standard; Protein; 359 AA.
XX
AC   R88870;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227A).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein; dominant negative phenotype.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Ala"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12904.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to an
CC   Ala residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were found to be resistant to
CC   STYLCV, i.e. the mutation results in a dominant negative phenotype.
XX
SQ   Sequence   359 AA;


  Query Match             59.8%;  Score 216;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-19;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   7
R88871
ID   R88871 standard; Protein; 359 AA.
XX
AC   R88871;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227H).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by His"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12905.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to a
CC   His residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             59.8%;  Score 216;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-19;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :|||||||| ||:||| |:|:|| || :|| :|:|  |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   8
```

```
R88872
ID   R88872 standard; Protein; 359 AA.
XX
AC   R88872;
XX
DT   07-NOV-1996  (first entry)
XX
DE   Sardinian tomato yellow leaf curl virus mutated Rep protein (K227R).
XX
KW   Nucleotide triphosphate binding site; DNA helicase; RNA helicase;
KW   modification; mutation; viral replication; deficient; inhibition;
KW   viral resistance; geminivirus; tomato yellow leaf curl virus;
KW   Sardinian isolate; STYLCV; transgenic plant; P-loop; C1 protein;
KW   AL1 protein.
XX
OS   Sardinian tomato yellow leaf curl virus.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 227
FT                   /note= "wild-type Lys has been replaced by Arg"
XX
PN   WO9608573-A1.
XX
PD   21-MAR-1996.
XX
PF   15-SEP-1995;   95WO-FR01192.
XX
PR   15-SEP-1994;   94FR-0011040.
XX
PA   (CNRS ) CENT NAT RECH SCI.
XX
PI   Gronenborn B;
XX
DR   WPI; 1996-179947/18.
DR   N-PSDB; T12906.
XX
PT   Prodn. of virus-resistant transgenic plants - using mutated genomic
PT   sequence from phytopathogenic DNA virus
XX
PS   Disclosure; Fig 13; 93pp; French.
XX
CC   Mutation of consensus amino acids in the NTP-binding site of
CC   geminivirus Rep protein is used to produce replication deficient
CC   viruses. The mutated viral nucleic acid is used for producing
CC   transgenic plants that are resistant to, or tolerant of, the native
CC   virus. The present sequence is a mutant form of the Rep (or C1)
CC   protein from the Sardinian isolate of tomato yellow leaf curl virus
CC   (STYLCV) in which the wild-type Lys227 residue has been changed to
CC   an Arg residue; transgenic Nicotiana benthamiana plants generated by
CC   transformation with the mutated virus were not resistant to STYLCV.
CC   In contrast, plants transformed with a virus in which Lys227 had been
CC   replaced by Ala were found to be resistant.
XX
SQ   Sequence   359 AA;


  Query Match             59.8%;  Score 216;  DB 17;  Length 359;
  Best Local Similarity   56.5%;  Pred. No. 3.1e-19;
  Matches   39;  Conservative   11;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       2 LVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLD 61
           | ||   :||||||||  ||:||| |:|:|| || :|| :|:|   |  |:  |||:|||||
Db     111 lewgtfqidgrsarggqqtandayakainagsksqaldvikelaprdyvlhfhninsnld 170

Qy      62 RIFDKTPEP 70
           ::|   | |
Db     171 kvfqvppap 179


RESULT   9
P70407
ID   P70407 standard; Protein; 353 AA.
XX
AC   P70407;
XX
DT   02-MAY-1991  (first entry)
XX
DE   ORF 4 gene product of Bean Golden Mossaic virus.
XX
KW   Plant vector.
XX
OS   Bean golden mosaic virus.
XX
PN   JP61257186-A.
XX
PD   14-NOV-1986.
XX
PF   10-MAY-1985;   85JP-0098108.
XX
PR   10-MAY-1985;   85JP-0098108.
XX
PA   (TEIJ ) TEIJIN KK.
XX
DR   WPI; 1987-159662/23.
DR   N-PSDB; N70630.
XX
PT   New DNA and hybrid DNA - used for recombinant vector of plants.
XX
PS   Disclosure; Fig 6; 24pp; Japanese.
XX
CC   The sequence encoding this protein may be taken along with the -a
CC   DNA sequence and a suitable resistance gene, and used to create a
CC   recombinant plant vector.
CC   See also N70629.
XX
SQ   Sequence   353 AA;


  Query Match             59.3%;  Score 214;  DB 8;  Length 353;
  Best Local Similarity   58.6%;  Pred. No. 5.4e-19;
  Matches   41;  Conservative   13;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  |||||||||| |::||: |:|||| | | || |::|: |: |: | ||: |||
Db     110 tiewgqfqvdgrsarggqqsandsyakalnadsiesaltilkeeqpkdyvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 eriffkvpep 179


RESULT  10
W34336
ID   W34336 standard; Protein; 361 AA.
XX
AC   W34336;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 protein.
XX
KW   Geminivirus; ToMoV; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
```

```
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93294.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.3; Page 57-58; 132pp; English.
XX
CC   This protein comprises the wild-type AC1 protein of tomato mottle
CC   virus (ToMoV), a geminivirus that has a bipartite genome.  The AC1
CC   gene (see T93294) must be expressed for efficient replication of
CC   the two genomic components, DNA-A and DNA-B.  The AC1 protein has a
CC   DNA binding site specific to the DNA-A common region, a DNA nicking
CC   activity, and an NTP binding activity.  The invention involves
CC   production of transgenic plants containing DNA comprising AC1 or C1
CC   wild-type or mutant sequences that negatively interfere in trans
CC   with geminiviral replication during infection.  Such transgenic
CC   plants are resistant to viral infection.  The AC1/C1 genes are
CC   especially from ToMoV, tomato yellow leaf curl virus or bean golden
CC   mosaic geminivirus (see T93282-93) and encode polypeptides (see
CC   W34324-35) that have mutations in the highly conserved DNA-nicking
CC   domain and/or the NTP-binding domain.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.5e-19;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :||||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  11
W34324
ID   W34324 standard; Protein; 361 AA.
XX
AC   W34324;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm.
XX
KW   Geminivirus; ToMoV-AC1dlm; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93282.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes  - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.4; Page 60-62; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336).  It is encoded by a mutant AC1 gene (see T93282) of
CC   tomato mottle virus virus (ToMoV), and carries 3 mutations in its
CC   NTP-binding domains.  The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome.  The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection.  The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.5e-19;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :||||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  12
W34325
ID   W34325 standard; Protein; 361 AA.
XX
AC   W34325;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm1.
XX
KW   Geminivirus; ToMoV-AC1dlm1; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93283.
```

```
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.5; Page 64-65; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm1, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336). It is encoded by a mutant AC1 gene (see T93283) of
CC   tomato mottle virus virus (ToMoV), and carries a mutation in an
CC   NTP-binding domain. The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection. The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.5e-19;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :||||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  13
W34326
ID   W34326 standard; Protein; 361 AA.
XX
AC   W34326;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato mottle virus AC1 mutant ToMV-AC1dlm23.
XX
KW   Geminivirus; ToMoV-AC1dlm23; AC1 gene; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato mottle virus isolate Florida.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93284.END.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT   golden mosaic geminivirus
XX
PS   Example 3.6; Page 67-69; 132pp; English.
XX
CC   This protein comprises a transdominant lethal mutant, designated
CC   ToMoV-AC1dlm23, of tomato mottle virus (ToMoV) AC1 protein (see
CC   W34336). It is encoded by a mutant AC1 gene (see T93284) of
CC   tomato mottle virus virus (ToMoV), and carries 2 mutations in an
CC   NTP-binding domain. The AC1 gene (see also T93294) must be
CC   expressed for efficient replication of the two genomic components,
CC   DNA-A and DNA-B, of the bipartite ToMoV genome. The invention
CC   involves production of transgenic plants containing DNA comprising
CC   geminivirus AC1 or C1 wild-type or mutant sequences that negatively
CC   interfere in trans with geminiviral replication during infection.
CC   Such transgenic plants are resistant to viral infection. The
CC   AC1/C1 genes are especially from ToMoV, tomato yellow leaf curl
CC   virus or bean golden mosaic geminivirus (see T93282-93) and encode
CC   polypeptides (see W34324-35) that have mutations in the highly
CC   conserved DNA-nicking and/or NTP-binding domains.
XX
SQ   Sequence   361 AA;


  Query Match             59.3%;  Score 214;  DB 18;  Length 361;
  Best Local Similarity   55.7%;  Pred. No. 5.5e-19;
  Matches   39;  Conservative   16;  Mismatches   15;  Indels    0;  Gaps    0;

Qy       1 TLVWGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNL 60
           |: ||:  :||||||||| |::||: |:|||||| : || ::||: |: :: | ||: |||
Db     110 tiewgdfqidgrsarggqqsandsyakalnassvqsalavlreeqpkdfvlqnhnirsnl 169

Qy      61 DRIFDKTPEP 70
           :||| | |||
Db     170 erifakapep 179


RESULT  14
W34329
ID   W34329 standard; Protein; 357 AA.
XX
AC   W34329;
XX
DT   27-APR-1998  (first entry)
XX
DE   Tomato yellow leaf curl virus mutant C1 protein pTYK104R #1.
XX
KW   Geminivirus; pTYK104R #1; C1 protein; transdominant mutation;
KW   transgenic plant; disease resistance.
XX
OS   Tomato yellow leaf curl virus strain Israel.
OS   Synthetic.
XX
PN   WO9739110-A1.
XX
PD   23-OCT-1997.
XX
PF   15-APR-1997;   97WO-US06300.
XX
PR   16-APR-1996;   96US-0015517.
XX
PA   (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA   (WISC ) WISCONSIN ALUMNI RES FOUND.
XX
PI   Ahlquist PG,  Hanson SF,  Luu HT,  Maxwell DP,  Stout JT;
XX
DR   WPI; 1997-526447/48.
DR   N-PSDB; T93287.
XX
PT   Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT   mutant genes - have increased resistance to geminivirus infection
PT   e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
```

PT golden mosaic geminivirus
XX
PS Example 4.5; Page 84-86; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the tomato
CC yellow leaf curl virus (TYLCV) C1 protein (see W34337) that is
CC required for replication. It is encoded by construct pTYK104R #1
CC (see T93287) of an infectious clone of TYLCV and carries a single
CC mutation in the DNA-nicking domain. The invention involves
CC production of transgenic plants containing DNA comprising
CC geminivirus C1 or AC1 wild-type or mutant sequences that negatively
CC interfere in trans with geminiviral replication during infection.
CC Such transgenic plants are resistant to viral infection. The
CC AC1/C1 genes are especially from TYLCV, tomato mottle virus or
CC bean golden mosaic geminivirus (see T93282-93) and encode
CC polypeptides (see W34324-35) that have mutations in the highly
CC conserved DNA-nicking and/or NTP-binding domains.
XX
SQ Sequence 357 AA;

Query Match 58.7%; Score 212; DB 18; Length 357;
Best Local Similarity 65.6%; Pred. No. 9.7e-19;
Matches 40; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

Qy 4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
:| : :|||||||| |::||| ||||: || ||| |::|| |: |: |||||:||||||
Db 111 fgvsqidgrsarggqqsandayaealnsgsksealnilkekapkdyilqfhnlssnldri 170

Qy 64 F 64
|
Db 171 f 171

RESULT 15
W34330
ID W34330 standard; Protein; 357 AA.
XX
AC W34330;
XX
DT 27-APR-1998 (first entry)
XX
DE Tomato yellow leaf curl virus mutant C1 protein pTYK225A #4.
XX
KW Geminivirus; pTYK225A #4; C1 protein; transdominant mutation;
KW transgenic plant; disease resistance.
XX
OS Tomato yellow leaf curl virus strain Israel.
OS Synthetic.
XX
PN WO9739110-A1.
XX
PD 23-OCT-1997.
XX
PF 15-APR-1997; 97WO-US06300.
XX
PR 16-APR-1996; 96US-0015517.
XX
PA (SEMI-) SEMINIS VEGETABLE SEEDS INC.
PA (WISC) WISCONSIN ALUMNI RES FOUND.
XX
PI Ahlquist PG, Hanson SF, Luu HT, Maxwell DP, Stout JT;
XX
DR WPI; 1997-526447/48.
DR N-PSDB; T93288.
XX
PT Transgenic plants expressing geminivirus AC1 and C1 wild-type and
PT mutant genes - have increased resistance to geminivirus infection
PT e.g. tomato mottle virus, tomato yellow leaf curl virus or bean
PT golden mosaic geminivirus
XX
PS Example 4.5; Page 88-90; 132pp; English.
XX
CC This protein comprises a transdominant lethal mutant of the tomato
CC yellow leaf curl virus (TYLCV) C1 protein (see W34337) that is
CC required for replication. It is encoded by construct pTYK225A #4
CC (see T93288) of an infectious clone of TYLCV and carries a single
CC mutation in the NTP-binding domain. The invention involves
CC production of transgenic plants containing DNA comprising
CC geminivirus C1 or AC1 wild-type or mutant sequences that negatively
CC interfere in trans with geminiviral replication during infection.
CC Such transgenic plants are resistant to viral infection. The
CC AC1/C1 genes are especially from TYLCV, tomato mottle virus or
CC bean golden mosaic geminivirus (see T93282-93) and encode
CC polypeptides (see W34324-35) that have mutations in the highly
CC conserved DNA-nicking and/or NTP-binding domains.
XX
SQ Sequence 357 AA;

Query Match 58.7%; Score 212; DB 18; Length 357;
Best Local Similarity 65.6%; Pred. No. 9.7e-19;
Matches 40; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

Qy 4 WGEAAVDGRSARGGCQTSNDAAAEALNASSKEEALQIIREKIPEKYLFQFHNLNSNLDRI 63
:| : :|||||||| |::||| ||||: || ||| |::|| |: |: |||||:||||||
Db 111 fgvsqidgrsarggqqsandayaealnsgsksealnilkekapkdyilqfhnlssnldri 170

Qy 64 F 64
|
Db 171 f 171

Search completed: February 3, 2001, 02:15:29
Job time: 3436 sec